PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/53, 33/566, 33/543, C07K 1/00, 2/00

A1

(11) International Publication Number: WO 98/58256

(43) International Publication Date: 23 December 1998 (23.12.98)

(21) International Application Number: PCT/US98/12580

(22) International Filing Date: 15 June 1998 (15.06.98)

(30) Priority Data:
60/049,634 16 June 1997 (16.06.97) US

(71) Applicant *(for all designated States except US)*: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; 308 Bynum Hall, CB#4105, Chapel Hill, NC 27599–4105 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: COLE, Ryszard [US/US]; 203 Longwood Drive, Chapel Hill, NC 27514 (US). LIU, Weiguo [CN/US]; 1647 Woodduck Lane, Wheeling, IL 60090 (US).

(74) Agents: SIBLEY, Kenneth, D. et al.; Myers, Bigel, Sibley & Sajovec, P.A., P.O. Box 37428, Raleigh, NC 27627 (US).

(81) Designated States: AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PEPTIDO OLIGONUCLEOTIDES (PONs) AND THEIR COMBINATORIAL LIBRARIES

(57) Abstract

The present invention provides for libraries of nucleotide–like substances reffered to as Peptido Oligonucleotides (PONs). The PONs of this invention consist of natural and unnatural D or L amino acids, purine or pyrimidine derived nucleobases and a four–carbon–chain connecting the nucleabases and the amino acids together through amide linkages to form a peptide backbone.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Title of the Invention

PEPTIDO OLIGONUCLEOTIDES (PONs) AND THEIR COMBINATORIAL LIBRARIES

Abstract

The present invention provides a novel embodiment or libraries enclosing large numbers of nucleotide like substances referred to as Peptido Oligonucleotides (PONs), and a powerful technique that efficiently select individual PONs against specific DNA or RNA targets in cell-lines for antisense therapeutics. The peptido oligonucleotides (PONs) in this invention consists of natural and unnatural L- or D-amino acids, purine and pyrimidine derived nucleobases, and a four-carbon-chain connecting the nucleobases and the amino acids together through amide linkages to form a peptide backbone. These three types of building blocks are arranged to allow a three-bond distance between the nucleobases and the backbone, and a six-bond distance between each nucleobase attached on the backbone. The arrangement provides the new PONs with optimum affinity to the complementary sequences of natural DNA or RNA molecules and by doing so render these analogs desirable features as potential antisense therapeutics. More importantly, this new construction of peptido oligonucleotides allows easy incorporation of various of functionalities in the molecule for a given sequence. By simply varying the connecting amino acids during standard peptide synthesis, one has the opportunity of generating very large numbers of antisense PONs that have different chemical and physical properties but are all complementary to a single target sequence. Against a defined 10 bp sequence of a target DNA or RNA, for example, a library consisting of 20^{10} PONs can be generated in theory by randomly choosing the connecting amino acids only from the proteogenic amino acid pool. The so obtained PON library is screened in a proper cell line bearing the target nucleic acid sequence, and only those PONs that efficiently penetrated the cell membranes, survived cellular degradation, and bonded strongly and selectively to the target sequences are selected. The connecting amino acid sequences of selected PONs are determined and sufficient quantities of the compounds are synthesized for further advanced testing. This new technology significantly improves the odds of developing clinically useful antisense therapeutics.

Field of the Invention

The invention generally relates to the creation and application of a large body of synthetic organic compounds that are capable of recognizing and binding to nucleic acids in a sequence-specific manner. Specifically, the invention provides a novel methodology for generation of very large numbers of defined mixtures of nucleotide like substances, i.e. peptido oligonucleotide (PON) combinatorial libraries, and the screening of the same for antisense

agents that are effective *in vivo* against specific DNA or RNA target sequences. The peptido oligonucleotide of this invention involves a hybrid of peptides and nucleotides with amino acids and nucleoside analogs alternately inter-connected through amide linkages to form oligomers that resemble nucleic acids. The geometry and topology of the base portions of the oligomer are preserved to function just like a nucleic acid in recognizing and base pairing with complementary sequences. The invention also relates to the novel synthetic processes for preparation of optically active amino acid nucleosides as building blocks and the construction and screening of peptido oligonucleotide combinatorial libraries.

Background

Antisense oligodeoxyribonucleotides (ODNs) have been offered as a major class of compounds for rational drug design (see Crooke, S. T. *Med. Res. Rev.*, 1996, 16, 319-344; Wagner, R. W., *Nature Med.*, 1995, 1:1116-1118; Milligan, J. F., Matteucci M. D., & Martin, J. C., *J. Med. Chem.*, 1993, 36, 1923-1937) These synthetic oligonucleotides can bind specifically by Watson-Crick base pairing to complementary DNA or RNA sequences and thus inhibit gene expression either by direct intervention of translation or transcription, or *via* activation of RNase H. Since, statistically, the base sequence of a 17-mer oligonucleotide occurs just once in the sequence of an entire human gnome, the selectivity of intervention with antisense ODNs of this length is very high. It is, therefore, possible to determine directly the chemical formula of a drug for treatment of a specific disease corresponding to the base sequence of the gene that causes the disease (see Wagner, R. W., *Nature*, 1994, 333-335; Uhlmann, E. & Peyman, A., Chem. Rev., 1990, 90, 544-584; Helene, C. & Toulme, J.-J., Biochim. Biophys. Acta, 1990, 1049, 99-125).

Although there have been numerous studies of antisense ODNs as therapeutic agents and pharmaceuticals including several on going clinical trials for treatment of acute myologenous leukemia (Reynolds, T., *J. Natl. Cancer Inst.*, 1992, 84, 288), infection by human immunodeficiency virus-1 (Alper, *J. biotechnology*, 1993, 11, 1225) and cytomegalovirus (Maister, P, *Bioworld Today*, 1994, 5, 3), and asthma (Nyce, J. W & Metzger, W. J., *Nature* 1997, 385, 721), there remain serious hurdles and barriers that need to be overcome (see Stein, C. A. & Cheng, Y. -C., Science, 1993, 261, 1004-1012; Stein, C. A., *Nature Med.* 1995, 1:1119-1121; Bennet, F. C., Chiang, M. Y., Chan, H. C., Shoemaker, J. E. E., Mirabelli, E. K., *Mol Pharm.*, 1992, 41, 1023-1033; Wagner, R. W., Matteucci, M. D., Lewis, J. L., Gutierrez, A. J., Moulds, C., Froehler, B. C., *Science*, 1993, 260, 1510-1513). Among these are the currently available ODNs' inability to cross the cellular membrane to reach the cytoplasm or nucleus, and their unstability toward degradation by endogenous nucleases. It was demonstrated that more than 70% of phosphorodiester ODNs

degraded in less than one hour incubation with cells (see Woolf, T. M., Jennings, E. G. B., Rebagliati, M. , Melton, D. A., *Nucleic Acid Res.*, 1990, 18, 1763-1769; Cohen, J. S., Ed. *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*; CRC Press: Boca Raton, FL, 1987; Chiang, M. Y., Chan, H., Zounes, M. A., Freier, S. M., Lima, W. F., Bennett, C. F., *J. Biol. Chem.*, 1991, 266, 18162-18171). The synthetic phosphorothioates and methylphosphonate ODNs are more resistant to nucleases but still have the problem of not efficiently penetrating the cell membranes. Besides, these two classes of ODNs are both produced as mixtures of diastereomers, and this could be the cause of some of the non-sequence-specific side effects observed (see Kibler-Herzog, L., Zon, G., Uznanski, B., Whittier, G., Wilson, W. D., *Nucleic Acid Res.* 1991, 19, 2979-2986; Lesnikowski, Z., J., Jaworska, M., Stec, W. J., *Nucleic Acid Res.* 1990, 18, 2109-2115). Currently, there is a very limited number of alternative antisense oligonucleotides available. Nielsen *at al*, U.S. Pat. No. 5,539,082, describes the synthesis and application of peptide nucleic acids (PNAs) which have peptide linkages replacing phosphodiester bonds and deoxyribose as the nucleotides' backbone and bind strongly to complementary RNA and DNA sequences. However, these oligomers are incapable of actively penetrating cell membranes and have to be delivered to their target by specialized technologies (Knudsen, H. &; Nielsen, P. E., *Nucleic Acids Res.* 1996, 24, 494-500). Although certain structure modifications have been conducted by attaching modifying groups to both N and C terminals of PNA molecules (U.S. Pat. No. 5, 539, 083), the construction and composition of PNA is generally inflexible to structure variations. Similar compositions of peptide-based oligonucleotides are described in PCT Int. Patent Publication WO 95/11909 and WO 95/04000. The former depicted a tetramer consisting threoninenucleosides which is difficult to synthesize for preparative purposes, and the latter deals with undefined mixtures of stereoisomers that severely limits their practical applications and the interpretation of.results. Besides their individual short-comings, a common disadvantage of the above antisense oligonucleotides is that for a given sequence of a target nucleic acid, only one complementary antisense oligo from each of the above category can be prepared. These oligos are then tested individually for antisense activities and sometimes modified with peripheral attachment of functionalities. Such a process is time-consuming and has been deemed inefficient in selecting effective antisense therapeutics. The complexity of interactions between antisense agents and biological systems should be considered when designing new antisense oligonucleotides. There are clearly many other factors besides Watson-Crick base pairing that decide the effectiveness of an antisense agent in inhibiting gene expression. For an antisense oligonucleotide to be effective as an therapeutic agent, it must meet the criteria of (1) can be synthesized easily and in bulk; (2) being stable *in vivo*; (3) being able to enter the target cell; (4) can be retained by the target cell; (5) being able to interact with

cellular targets: and (6) not interact in a non-sequence-specific manner with other macromolecules. None of the currently available ODN analogs meet all of these criteria (Stein, C. A. & Cheng, Y.-C., *Science* ,1993, 261, 1004-1012). A practical approach for selecting such an antisense oligonucleotide would be that for a given target sequence of a gene, a large number of oligonucleotides with the same antisense sequence but different structures and chemical/physical properties are synthesized simultaneously and then tested for the desired biological activities as a group in order to identify a proper candidate for further development. In a conventional drug development program, a lead compound is generated first for a given therapeutic target. This initial drug candidate may have some desired biological activities along with undesired side effects or toxicity. Based on the structure of this lead compound, an array of derivatives and analogs are prepared and tested. From them, the final drug with the highest biological activity and lowest side effects is identified. In this invention, the conventional practice in developing antisense drugs by generating one antisense molecule for each target gene sequence is replaced with the generation of a large group of antisense molecules for each target gene sequence. The antisense sequence of the oligonucleotides is conserved as the pharmacophore as in the lead compound, while the rest of the structure such as the types of backbone, the various side chains on the backbone, and the functional groups on the bases are altered to generate a great variety of analogs and derivatives from which the best antisense molecule for the purpose is selected. There are so far no other alternatives for convenient generation of antisense combinatorial libraries against a single target sequence. The technology described in this invention has a clear advantage in the rapid development of very large numbers of new, potent antisense agents which have the required properties of stability, affinity, permeation, and ultimately, favorable pharmacokinetics. That is indeed the objective of this invention.

Objectives of the Invention

The primary objective of the present invention is to provide a new class of polymeric molecules capable of forming duplex or triplex structures with nucleic acids in a sequence specific manner.

A further objective of the invention is to formulate and construct analogs that function like Peptide Nucleic Acids in the existing art but overcome some of the disadvantages associated with PNA such as solubility and cellular uptake.

Another objective of the invention is to provide methods of generating combinatorial libraries of defined sequence oligonucleotide analogs which include large numbers of antisense molecules that have a same base sequence but different chemical, physical, and biological properties.

Yet, another objective of the invention is to provide a methodology of effectively selecting desired antisense agents from libraries consisting of vast numbers of oligonucleotide analogs bearing same base sequences but different functionalities.

Summary of the Invention

The present invention provides a novel embodiment or libraries of large numbers of nucleotide like substances referred to as Peptido Oligonucleotides (PONs), and a powerful technique that efficiently select individual PONs against specific DNA or RNA targets in cell lines for antisense therapeutics. The peptido oligonucleotides (PONs) in this invention consists of natural and unnatural L- or D-amino acids, purine and pyrimidine derived nucleobases, and a four-carbon-chain connecting the nucleobases and the amino acids together through amide linkages to form a peptide backbone. These three types of building blocks are arranged to allow a three-bond distance between the nucleobases and the backbone, and a six-bond distance between each nucleobase attached on the backbone. The arrangement provides the new PONs with optimum affinity to the complementary sequences of natural DNA or RNA molecules and by doing so render these analogs desirable features as potential antisense therapeutics. What is more important, this new construction of peptido oligonucleotides allow easy incorporation of various of functionalities in the molecule for a given sequence. By simply varying the connecting amino acids during standard peptide synthesis, one has the opportunity of generating very large numbers of antisense PONs that have different chemical and physical properties but are all complementary to a single target sequence. Against a defined 10 bp sequence of a target DNA or RNA, for example, a library consisting of 20^{10} PONs can be generated in theory by randomly choosing the connecting amino acids only from the proteogenic amino acid pool. The so obtained PON library is screened in a proper cell line bearing the target nucleic acid sequence, and only those PONs that efficiently penetrated the cell membranes, survived cellular degradation, and bonded strongly and selectively to the target sequences are selected. The connecting amino acid sequences of selected PONs are determined and sufficient quantities of the compounds are synthesized for further advanced testing. This new technology will significantly improve the odds of developing clinically useful antisense therapeutics.

At least a portion of the PONs of the invention has the stereochemically defined composition in the form of $S\text{-}(pX\text{-}AA)_n\text{-}Y$

Wherein:

S is a hydrogen or a linker or a modifying group or a peptide.

Y is a hydrogen or a modifying group or an amino acid or a peptide.

AA is one of any natural and unnatural amino acids excluding pX.

pX is an optically active amino acid nucleoside having the structure of

$HOOCCHNH_2CH_2CH_2$-X

where X is any one of the nucleobase or their derivatives including thymine. cytosine, uracil, adenine, and guanine.

n = 1 or more (e.g., 1 or 2 to 20, 30, 50 or 100).

The composition involves defined chiral centers bearing either (R) or (S) configurations.

The compounds of the invention generally are prepared by solid phase peptide synthesis techniques. A novel enzyme catalyzed enantioselective hydrolysis reaction was applied to prepare both (R) and (S) enantiomers of the pX through resolution of the racemic mixtures synthesized by established methods.

Brief Description of the Figures

Fig. 1 is the synthetic scheme for preparation of γ-bromo-α-aminobutyric acid derivatives.

Fig. 2 is the synthetic scheme for preparation PON monomer pT ,where p is α-aminobutyric acid, and T is thymine.

Fig. 3 is the synthetic scheme for preparation PON monomer pC, where p is α-aminobutyric acid, and C is cytosine.

Fig. 4 is the synthetic scheme for preparation PON monomers pG and pA, where p is α-aminobutyric acid, G is Guanine and A is adenine.

Fig. 5 is the reaction scheme for enzymatic resolution of racemic ethyl α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate ((dl)-Boc-pT-OEt) to obtain both (S)-, and (R)-Boc-pT.

Fig. 6 is the chrial HPLC monitoring of the resolution process. The 4 chromatograms are (dl)-Boc-pT-OEt, its reaction mixture with papain, (S) and (R) Boc-pT, and isolated (S)-Boc-pT.

Fig. 7 is the synthetic scheme for preparation of deoxythymidine-2'-amino-5'-carboxylic acid.

Fig. 8 is the synthetic scheme for preparation of dideoxycytodine-2'-amino-5'-carboxylic acid.

Fig. 9 is a demonstration of reactions catalyzed by uridine phosphorylase and purine nucleoside phosphorylase.

Fig. 10. is the general synthetic scheme for preparation of 2'-amino-5'-carboxylic acid of purine didexoyribonucleosides through base-change reactions catalyzed by a combination of by uridine phosphorylase and purine nucleoside phosphorylase.

Fig. 11. is a demonstrattion of properly protected amino acid nucleosides for peptide synthesis.

Fig. 12. is the synthetic scheme for preparation of properly protected dideoxycytodine-2'-amino-5'-carboxylic acid.

Fig. 13 is the chemical structure of the PON (lys-(pT-gly)$_{10}$-gly-NH_2).

Fig. 14 is a computer model of a double helex formed between a PON (ala-pX)$_n$ and a complementary single stranded DNA (dX)$_n$.

Fig. 15 is an example of constructing primary PON libraries applying one-bead-one-peptide approach.

Fig. 16 is an example of constructing secondary PON libraries by coupling mixtures of the connecting amino acids.

Detailed Description of the Invention

The term "oligonucleotides" as used in connection with this invention refers to polymeric molecules having repeated units formed in a specific sequence from naturally occurring bases and sugars joined together through phosphodiester bonds. These molecules include fragments of DNA, RNA, and their derivatives. The term oligonucleotide analogs refers to those compounds that function like oligonucleotides but have modified or completely re-designed structures. The term peptide nucleic acid (PNA) relates to a special group of oligonucleotide analogs having a peptide backbone with side chains having nucleobases that are capable of engaging in hydrogen bonding with an oligonucleotide having a complementary sequence. The peptido oligonucleotides (PONs) of the present invention refers to a new class of peptide nucleic acids through novel assembly of subunits consisting of natural and unnatural amino acids, natural and modified nucleobases, and a bridging molecule that effectively link the nucleobases with the amino acids. The bridging molecule in this invention itself is an amino acid in nature. It is attached to nucleobases through displacement of an ω-leaving group in a N-protected α-amino butyric acid by one of the nucleophilic nitrogen's on the nucleobase. The term "interconnecting amino acids" in this invention relates to a series of natural and unnatural amino acids including D-, and L-amino acids, α,α-disubstituted amino acids, and amino acids with secondary amino groups or amino groups incorporated in a ring system.

The term "complementary" indicates that a particular sequence of bases is able to pair with corresponding bases in a given target sequence either through Watson-Crick or Hoosteen base-pairing.

The term "combinatorial library" refers to an embodiment of usually a large number of different substances generated systematically and simultaneously through ordered, well controlled synthetic steps in combination with a random distribution of a number of defined building components.

The basic building units for the peptido oligonucleotides (PONs) of this invention include an array of α-aminobutyric acid derived amino-acid-nucleosides, and are referred to as PON monomers designated as pX, while X equals to A (adenine), T (thymine), C (cytosine), G (guanine), U (uracil), and their modified analogs; p represents the α-aminobutyric acid portion of the nucleoside serving as a spacer and a linker with free or protected amino carboxylic acid bifunctionalities. A general representation of pX is illustrated in the following structure: $HOOCCHNH_2CH_2CH_2$-X
wherein, X is a nucleoside base or its modified derivatives

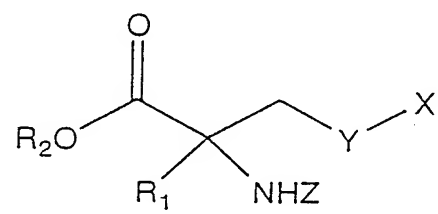
The general approach for preparing PON monomers is to attach the properly protected amino acid directly to various protected or unprotected nucleobases and then manipulate the adducts to release the desired functionalities. In this invention, we also disclose a novel chemo-enzymatic process for preparation of optically active 2-amino-4-acyl-butyric acid analogs (acyl = pyrimidine and purine nucleobases) as PON monomers or building blocks.

Syntheses of racemic 2-aminobutyric acid substituted with pyrimdiine and purine nucleobases at carbon-4 have been reported in the prior arts (Koch, T. & Buchardt, O., *Synthesis*, 1993, 1065; Nollet, A. J. H., Huting, C. M., Pandit, U. K., *Tetrahedron*, 1969, 25, 5971; ; Nollet, A. J. H., Huting, Pandit, U. K., *Tetrahedron*, 1969, 25, 5994) The process generally involve heating homoserine or homoserine-γ-lactone with hydrogen bromide in acetic acid to produce α-amino-γ-bromobutyric aicd hydrogenbromide (Fig.1). The bromoacid is treated with hydrogen chloride gas in anhydrous ethanol to produce the ethyl ester which is then protected at the amino group with Boc by reacting with di-tert-butyldicarbonate in aqueous sodium carbonate to yield ethyl α-amino-γ-bromo-N-*t*-butoxycarbonylbutyrate. The protected γ-bromo-aminoester is then attached to various nucleobases by displacement of the bromo group of the amino ester with the nitrogen's in the nucleobases (Fig.2-4). The resulting amino ester nucleoside is hydrolyzed by sodium hydroxide or lithium hydroxide to generate the free acid. This synthetic sequence involves very strong acid and high temperature conditions which could lead to racemization at the α-carbon if an optically active amino acid is involved. Indeed, we have observed repeatedly partial or complete racemization at the α-carbon during preparation of the optically active products with either L-, or D-homoserine as starting materials.

Unlike the process for preparation of an optically active pharmaceutical in which asymmetric synthesis of a single enantiomer is preferred over enantiomeric resolution where

half of the product has to be discarded, the synthesis of optically active PON monomers such as the 2-aminobutyric acid derivative is, at present, intended for both enantiomers. By simply resolving the racemic final product we could obtain both L-, and D-amino acid nucleosides for construction of a variety of homogenous PON stereomers. Libraries of PON stereomers can be screened for binding affinity with complementary nucleic acids and the optimum conformation and stereochemistry of the PON can be selected.

A part of the present invention refers to a process in which two enantiomers in a racemic mixture of amino acid nucleoside are separated physically after treating the mixture with an enantioselective reagent. This reagent reacts preferentially to one of the two enantiomers in the racemic mixture yielding a product with different chemical structure from the un-reacted enantiomer thereby generating the differences for the separation of the two. A general representation of the racemic amino nucleosides is illustrated in the following structure



wherein:

X is a nucleoside base or its modified derivatives

R_1 is H; alkyl or substituted alkyl; alkenyl or substituted alkynyl; alkaryl or substituted alkaryl; aralkyl or substituted arakyl; cyclic or heterocyclic ring systems.

R_2 is H; alkyl or substituted alkyl; alkenyl or substituted alkynyl; alkaryl or substituted alkaryl; aralkyl or substituted arakyl; alcyclic; cyclic and heterocyclic ring systems.

Y is CH_2 or CH_2CH_2 or O or S.

Z is a protecting group including Fmoc, Boc, Cbz, Pht, etc.

The enantioselective reagents include optically active molecules bearing acid or base functionalities that serve the purpose of general acid or base catalyses. Representing members of these reagents include hydrolytic enzymes such as papain, trypsin, subtlisine, chymotrypsin, acylases, esterases, lypases, and other proteases. The reaction involves enantioselective hydrolysis of either the carboxylic ester R_2 or the nitrogen protecting group Z. The resulting optically active carboxylic acid or the free amine are both significantly more water soluble than the un-reacted starting material and are thus easily separated. For example (Fig.5), the racemic ethyl α-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyrate is treated with papain in acetonitrile/water (2:8) at room temperature for

several hours. The progress of the reaction is monitored by HPLC (Fig.6). After about 50% ester to acid conversion, the reaction mixture is extracted repeatedly with hexane. The hexane extracts are combined, dried, and concentrated to dryness to obtain the optically pure ethyl α-(R)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyrate which is then hydrolyzed chemically with NaOH to give α-(R)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid . The aqueous layer is concentrated under vacuum to a small volume and then 100% ethanol is added to precipitate the enzyme protein and inorganic salt. After filtration, the clear filtrate is further concentrated to near dryness followed by crystallization to give α-(S)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid in 97% ee. The same resolution process have also been applied to the preparation of both enantiomers of α-amino-γ-(1-(4-amino-2-hydroxypyrimidyl)-N-*t*-butoxycarbonylbutyric acid, α-amino-γ-(1-(2,4-dihydroxypyrimidyl)-N-*t*-butoxycarbonylbutyric, α-amino-γ-(7-(4-purinyl)-N-*t*-butoxycarbonylbutyric acid, and α-amino-γ-(7-(2-amino-4-dihydroxypurinyl)-N-*t*-butoxycarbonylbutyric acid. The analogues enzymatic resolution processes were also carried out in reverse reactions in organic solvents starting from the free acids. For example, racemic α-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid is treated with papain in a mixture of ethanol/hexane/buffer (1/1/3) at room temperature for 24 hours. After about 50% acid to ester conversion, the reaction mixture is filtered to remove the enzyme, diluted with water, and extracted repeatedly with hexane. The hexane extracts are combined, washed with water, dried, and concentrated to dryness to obtain the optically pure active α-(S)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyrate. The aqueous layer is concentrated under vacuum to near dryness followed by crystallization to give optically pure α-(R)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid. All these optically active amino acid nucleosides are used as the basic building blocks - PON monomers (pXs), for the construction of a vast variety of PONs and PON libraries.

Besides α-aminobutyric acid derived nucleosides, other nucleoside analogs bearing aminocarboxylic acid functionalities can also serve as building blocks for construction of PONs. These alternative PON monomers can be prepared ether by attaching amino acids to the nucleobases or by direct derivatization of deoxyribonucleosides. To introduce amino acid bifunctionality into dexoyribonucleosides. for example, the commercially available 2',3'-

dideoxy-3'-azido-thymidine (AZT) (Fig.7) is used as one of the starting materials. AZT is first oxidized to the 5'-carboxy-derivative by an appropriate oxidizing agent such as chromic acid, potassium permagnate, and ruthenium trichloride. Since direct hydrogenation of the acid resulted to its decomposition, the methyl ester derivative is prepared and reduced by hydrogenation on pd/C to give the amino ester which is then hydrolyzed by sodium ethoxide in ethanol/water to yield the deoxythymidine amino acid.

For synthesis of the corresponding amino acid of cytodine (Fig.8), deoxyuridine was treated with diphenylsulfide in DMF to yield the 3',2-dehydrouridine which is oxidized to the 5'-carboxylic acid and then converted to the ethyl ester. Nucleophilic attack of the ester by sodium azide in DMF catalyzed by lithium sulfate give the 3'-azide-5'-carboxydeoxyuridine. Treatment of the uridine derivative with phosphorus oxychloride and triozole in pyridine followed by amination give the deoxycytodine azide. During amination, the 5'-carboxylic ester was hydrolyzed to the free acid which was re-esterified followed by Pd/C catalyzed hydrogenation and base hydrolysis to yield the desired deoxycytodine amino acid.

The corresponding amino acids of the purine deoxyribonucleosides are prepared through a novel process based on base exchange reactions catalyzed by a group of enzymes termed Nucleoside Phosphoralases (Fig.9). Uridine phosphorylase catalyzes the equilibrium reactions of uridine and other pyrimidine nucleoside with inorganic phosphate to give the free pyrimidine nucleobase and ribose-1-phosphate. Purine nucleoside phosphorylase catalyzes the same equilibrium reactions between a purine-nucleoside and its free base plus ribose-1-phosphate, in which the equilibrium is heavily tilted toward nucleoside formation. Since ribose-1-phosphate is the common intermediate in both enzyme reactions, when the two reactions are combined the ribose-1-phosphate generated from the first reaction is immediately utilized in the second reaction to react with purine nucleobases in forming purine nucleoside. Due to the depletion of ribose-1-phosphate, more pyrimidine nucleoside is converted to pyrimidine base and ribose-1-phosphate which is further used in the second reaction to make purine nucleoside. When excess amount of the purine nucleobase is added, the conversion from pyrimidine nucleoside to its purine analog can be driven to 80-90% completion. Both enzymes accept a very broad range of substrates, require no co-factors, and can be produced in large quantities. In actual industrial applications, the two enzymes are co-immobilized in agarose beads, and suspended in phosphate buffer containing the substrates of pyrimidine nucleoside and purine bases. The mixture are stirred at 30 °C until more than 80% of starting material is converted to the corresponding purine nucleoside, and then filtered to recover the enzyme. The filtrate is first extracted with chloroform to remove the nucleobases, and then with ethyl acetate or butanol to obtain the product. Using 2',3'-dideoxy-3'-amino-5'-carboxythymidine prepared from AZT as the starting material to exchange base with a purine

analog such as adenine or guanine catalyzed by the combined nucleoside phosphorylases system as above, the corresponding purine nucleoside amino acid are prepared in reasonable yields (Fig.10). These amino acid nucleosides are then properly protected and subjected to peptide synthesis for PON construction.

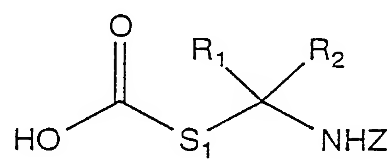
A prototype of PON containing $HOOCCHNH_2CH_2CH_2$-thymine as pX (pT) and glycine as the connecting AA was assembled by standard solid phase peptide synthesis. Boc-Gly-MBHA (p-methylbenzhydrylamine) resin was used as starting material on an automated peptidē synthesizer. Coupling of optically pure (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric to the deprotected glycine-resin using 2-(1H-benzotriozole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) pT-glycine-resin. After deprotection, Boc-glycine was coupled to the peptide chain followed by (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric again. The cycle was repeated until the 10th (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric was attached to the growing peptide chain. The resin was then deprotected and coupled with L-lysine as the N-terminal residue to increase the water solubility of the resulting PON molecule. The peptide was then cleaved from the MBHA resin and purified to obtain the target PON with lysine at the N terminal and glycine amide at the C-terminal as lys-(pT-gly)$_{10}$-gly-NH_2 (p = α-aminobutyric acid) (Fig.13).

The ability of the PON to form a duplex with single stranded DNA or RNA was studied by molecular modeling (Fig.14). A standard double-stranded DNA was called on from the data base. The phosphorodiester linkages on one strand of the DNA was replaced with the (S)-alanine- (S)–α-aminobutyric acid linkage as in the PON molecule. The α-carbons and α-nitrogen's in both alanine and α-aminobutyric acid were kept in the same planes as their neighboring carbonyl groups. Rotating the two planes along Y-axis yields several conformations of lower energy than that of the original phosphodiester backbone. More importantly, in most of these lower energy conformations, the side-chain (Me) of alanine points away from the double-strand and the staking nucleobases, suggesting that substitution on connecting AAs of a PON is unlikely to hinder its base-pairing with the complementary nucleic acid.

Thermal stability's of the duplex of lys-(pT-gly)$_{10}$-gly-NH_2/(dA)$_{12}$, lys-(pT-gly)$_{10}$-gly-NH_2/(dA)$_5$dT(dA)$_6$, and lys-(pT-gly)$_{10}$-gly-NH_2/(dA)$_5$(dT)$_2$(dA)$_5$ were tested by measuring the melting points of the hybrids on a Gilford Response apparatus following procedures described by Egholm *at. al.* (Egholm, M., Nielsen, P. E., Buchardt, O., Berg, R. H. *J. Am.*

Chem. Soc. 1992114, 9677-9678). The T_m of lys-(pT-gly)$_{10}$-gly-NH_2/(dA)$_{12}$ as a completely complementary PON-DNA duplex was recorded as 74 °C, very close to the T_m value (73 °C) for the corresponding PNA-DNA duplex but significantly higher than that (25 °C) of (dT)$_{12}$/(dA)$_{12}$ as a standard DNA-DNA duplex. Melting temperatures (T_m) for PON-DNA duplex with one and two mismatches are 59 °C and 49 °C respectively. These results further indicate that the PONs of the said composition have the ability to recognize and bind strongly and specifically to complementary nucleic acid sequences, and therefore, demonstrates their potential utilities as antisense agents for research, diagnostic, and therapeutic applications.

The peptido oligonucleotide (PON) combinatorial libraries of the present invention are generated by connecting PON monomers alternating with natural or synthetic amino acids as spacers. The PON monomers and the spacer amino acids are linked in such a way that the nucleobase sequence of the resulting PONs are well defined while the backbone that connects these nucleobases are altered in a combinatorial manor to generate large numbers of antisense molecules complementary to a given target sequence. The resulting peptido oligonucleotides produced by this process will possess different chemical, physical, and biological properties. A typical PON consists at least a portion of a simple repeat of pX and AA as core PON in the form of -(AA-pX-)$_n$-, where n is Ø or as above and AA represents natural and synthetic amino acids with a general structure of



wherein

R1 is H; alkyl or substituted alkyl; alkenyl or substituted alkynyl; alkaryl or substituted alkaryl; aralkyl or substituted arakyl; alcyclic; cyclic and heterocyclic ring systems.

R2 is H; alkyl or substituted alkyl; alkenyl or substituted alkynyl; alkaryl or substituted alkaryl; aralkyl or substituted arakyl; acyclic; cyclic and heterocyclic ring systems.

Z is H or alkyl or alkenyl or a protecting group including Fmoc, Boc, Cbz, Pht, etc.

S is a bond or an atom or a group of atoms

S, R1, R2 and Z can be interconnected in one or more ring systems.

pXs and AAs are coupled together through standard reactions for peptide bond formation including both solution and solid phase peptide synthesis. Generally, the PONs are assembled on solid phase resins following Merrifield method or its modified versions. For a specific PON library, the sequence of the PONs is predetermined based on the sequence of the

target gene fragment. These target sequences are usually selected according to their sensitivities to antisense inhibitions and the functions of their protein products. A target RNA sequence of AATTTCCGGG for example, dictate the types of pX and their order of introduction as pTpTpApApApGpGpCpCpC for the corresponding PON library. Each pX is separated by an amino acid (AA) as the spacer in the actual PON molecules for proper base pairing with the target. The general structure of core PONs for this target library is therefore pT-AA-pT-AA-pA-AA-pA-AA-pA-AA-pG-AA-pG-AA-pC-AA-pC-AA-pC-AA, where each and every individual AA can be any amino acids in any combinations. If AAs are only draw from a pool of the 20 gene encoded natural amino acids, the library will in theory contain 20^{10} members of PONs with the same pX sequence but different AA combinations.

The size of a certain PON library for a specific target sequence is defined by the maximum number of PONs it covers, which, in turn depends on the length of the target sequence and the variety of AAs incorporated. It is generally expressed as x^n, where n is the number of bases in the target sequence or in the PONs and x is the number of amino acids among which the AAs are selected. This expression only applies to libraries of core PON structures. The actual size of a PON library can be larger than x^n when extra AAs or a strings of AAs are attached to either C or N or both terminals of the core PONs.

A PON library for a target sequence can be further divided into sublibraries where the pX sequence is the same but the connecting AAs of the PONs are subdivided according to their chemical, physical and biological properties. A library with PONs containing partially or wholly D-amino acids as the connecting AAs is regarded as a D-AA sublibrary; while the one with PONs containing pXs bearing a D or R chiral center is referred as a D-pX sublibrary. If the PONs in a library contain both D-AAs and D-pXs, then the library is called a D-PON library assuming the unspecified PON libraries as L-PON sublibraries. In the same principle, a PON library with PONs containing, as the connecting AAs, predominantly lipophilic amino acids, or hydrophilic amino acids, or anionic amino acids, or cationic amino acids, or α,α-disubstituted amino acids is referred to as a lipophilic sublibrary, or a hydrophilic sublibrary; or an anionic sublibrary, or a cationic sublibrary, or a disubstituted sublibrary. A sublibrary can enclose sub-sublibraries according to the further differences of fuctionalities incorporated into the PONs. For example, a cationic PON sublibrary can contain sub-sublibraries with PONs incorporating predominantly lysine, or arginine, or other positively charged amino acids as the connecting AAs. Each sub-sublibrary can be divided further and further into branched smaller libraries based on more and more detailed differentiation's of the fuctionalities.

Careful planing, grouping and construction of appropriate combinations of PON sublibraries is essential for successful screening and selection of antisense PONs with desired

activity and properties. A target nucleic acid sequence, or an antisense PON against the sequence, typically contains 10 to 20 nucleobases. If the PON library against this sequence include all possible amino acids as the connecting AAs, it would be too large to be practically constructed due to the demand for huge amounts of the total mixture in order to include every possible PONs in a detectable quantity. Currently available screening and analytical techniques require a minimum of about 100 picomoles of a PON present in the mixture to be effectively selected. Assuming the average molecular weight is 100 for connecting AAs and 200 for pXs, the average molecular weight for a PON with 10 nucleobases is approximately 3000. If a specific 10 nucleobase PON library include all 20 natural amino acids independently as connecting AAs, the total number of individual PONs in the library will be 20^{10}. For each member of the PON in the library to be present in a quantity above 100 picomoles, the mass of the total mixture will have to be larger than 3.1 metric tons. Such a library is obviously impractical to construct and screen. It is therefore necessary to construct smaller sublibraries to probe certain general features of antisense PONs at the beginning of the research. Once some of these general features are understood, further branched sublibraries are constructed for another round of investigations until the desired PONs are selected.

Several tiers of parameter sets are measured for determination of structure activity relationships (SAR) of PONs based on sublibrary screening and testing. Some of the most important parameters of antisense PONs include specificity and affinity of binding to complementary sequences of nucleic acids, stability to intracellular degradation, solubility in aqueous media, and the ability to penetrate cell membranes to reach the nucleic acid targets. The specificity and affinity of binding to nucleic acids is the most critical measurement of any antisense molecules and is therefore the first set of criterion to be tested for construction and selection of PON sublibraries. As governed by the principles of Waterson-Crick Base Pairing, the ability of a PON to recognize and bind to complementary sequences of nucleic acids is primarily influenced by the space between the neighboring nucleobases in the PON molecule and by the distance from the nucleobase to the backbone. Stereochemistry is another first tier variable that affect binding. Also, electronic charge is an important factor since the target nucleic acids are highly charged molecules. Contributions from these first tier variables to the binding affinity of PONs are generally independent of other structural changes confined within the scope of their definitions. Therefore, the very first sets of PON sublibraries are constructed and screened primarily for their binding affinities to complementary DNA or RNA sequences. These sublibraries consist of PONs with fundamental structure differences in terms of stereochemistry, electronic charge, rotation flexibility, length of the spacer and the distance between neighboring nucleobases. Many of these primary structure features in a PON library are imposed by the pX portion of the molecules. When the X's are chosen only from natural

nucleobases, the connecting molecule p becomes the critical building unit. Although this invention deals primarily with PON libraries consisting α-aminobutyric acid-based nucleosides as pXs, other amino acid-nucleosides such as (α)-, and (β)-2'-amino-5'-carboxyl-deoxyribonucleosides, and D-, and L-threonine-based nucleosides are also included as alternative PON monomers. PONs consisting of these nucleosides could be ether invaribale or variable. Invariable PONs contain a single type of p as pX throughout a PON molecule resulting an even construction in terms of distances between each nucleobases and from the nucleobase to the backbone. The variable PONs incorporate more than one type of p within each PON chain so that the property and geometry of each pX in the PON molecule could be different. PON sub-libraries bearing primary structure features are screened, usually in vitro, for their affinities and specificity's of binding to target DNA or RNA sequences (Fig.15).

Once a proper type of construction is chosen for the PON from screening of first tier PON libraries, a set of second tier PON sub-libraries are generated by varying the connecting AAs in the PON molecule with fixed primary structure features (Fig.16). Since the changes of AAs in a PON molecule does not significantly affect the distances between each nucleobases and the distance from the nucleobase to the backbone, their effect on the binding affinities of the PON to the complementary DNA or RNA sequences is secondary to the changes of pXs, except the electronic charge. Positively or negatively charged functional groups on the backbone will likely impose certain effects on the binding affinity of the resulting PONs due to the ionic or electrostatic interactions between the PON and the nucleic acid target. But these effects are unlikely to be affected by changes of other connecting AAs as long as these changes do not create new charges. In most cases, changes of AAs in a PON molecule result in the changes of chemical and physical properties of the compounds such as polarity, solubility, lipophilicity, stability, antigenicity, etc. without much effects on binding affinity to the target sequence of nucleic acids. These changes are thus exploited for fine-turning the PON's chemical and physical properties to achieve desirable pharmacokinetic profiles in vivo. Among some of the most important goals to be achieved by fine-tuning the PON's secondary structures are to maximize the PON's ability to penetrate cell membranes and to minimize its non-specific interactions with other endogenous macromolecules. Through these consecutive two-stage construction and screening of primary and secondary PON libraries, an antisense PON that can be readily delivered in vivo to the target nucleic acid and binds to it with strong affinity and high specificity could be selected in relatively a short period of time, providing a powerful tool for generation of effective antisense therapeutics in the treatment of gene related diseases.

BINDING OF PONs TO TARGET NUCLEIC ACID AND PON LIBRARY SCREENING:

Homogeneous PONs refer to those peptido oligonucleotides having a uniformed linker molecule as p, such as 2-aminobutyric acid, and a single amino acid as the connecting AA. These PONs are prepared for representative sequences such as $(gly\text{-}pT)_{10}$, and are usually constructed on resin beads, released and isolated as free peptides, and tested individually for binding to complementary target nucleic acid sequences such as $(dA)_{10}$. Measurement of melting temperatures of the resulting duplex or triplex of PON-nucleic acid hybrids are the most commonly used methods for determining their binding affinity. The stronger the PON binds to the complementary nucleic acid, the higher the hydribs' melting temperature, at which half of the complex dissociates to two single strands. The ability of a PON to recognize and bind to a complementary nucleic acid sequence can also be tested by gel retardation experiments in which complementary and non-complementary sequences of single strand nucleic acids are exposed to the testing PON before subjecting to agarose gel electrophoresis. The RNA strands that are not complementary to the PON sequence migrates normally on the gel, but the ones that are complementary to the PON will move much slower on the gel due to the formation of PON-RNA hybrids.

Heterogeneous PONs have two or more amino acids serving as the connecting AAs and sometimes include different combinations of p build into pXs. These compounds are, in most cases, synthesized as defined mixtures either in a solution or a solid phase following standard combinatorial library generation methods. The solid phase PONs are constructed on synthetic resins according to established methods for preparation of peptide combinatorial libraries. Starting with Boc-amino acid resins for example, coupling the deprotected AA-resins with a properly protected pX followed by deprotection gives the resin-AA-pX. The nucleobase X is selected among A, T, C, G, and U depending on the sequence of the target nucleic acid, while p is the α-aminobutyric acid linker or other alternative spacer with aminocarboxylic acid bi-functionalities. For a PON library against a target sequence of $(dA)_{10}$, for example, the pX can be selected from at least the following three types of amino acid nucleosides and their stereoisomers: 3'-amino-5'-carboxythymidine (A), 3-oxy-(1-thymidyl)-threonine (B), and 4-(1-thymidyl)-2-aminobutyric acid (C). The AA-resins are evenly divided into three groups and coupled respectively with properly protected A, B, and C. After deprotection, the resins are combined, well mixed and coupled with a specific Boc-AA. After removing the Boc group, the resins are, again, evenly divided into three groups and coupled with A, B, and C respectively. This process is repeated until the resulting peptide chain contains 10 pX units. At this point, the library would have contained in theory a total of 59049 different types of PON molecules all having 10 thymine units complementary to $(dA)_{10}$ with one or more of defined amino acid as the connecting AA. Similar $(pT)_{10}$ libraries can be generated with a

single 4-(1-thymidyl)-2-aminobutyric acid as pX, while a variety of positively or negatively charged amino acid are incorporated into the PON chain as connecting AAs in a combinatorial fashion. These resin libraries are then screened for binding affinity with $(dA)_{10}$ by appropriate selection methods.

Screening of primary PON libraries are conducted mostly *in vitro*, and involves relatively smaller libraries. In general, these methods call for the suspension of the PON resin library in a proper buffer that contains a natural or synthetic oligonucleotide bearing the target sequence and labeled with fluorescent or chemical laminating groups. The mixture is warmed to about 90 °C and then cooled slowly to 4 °C with stirring. After filtration and wash, the resins are resuspended, respectively, in buffers of different temperature, such as 70, 80, 90 and 100 oC, stirred, filtered, washed, and visualized under UV light or with chemical treatment. Those few resins that still illuminate after high temperature wash are selected and decoded to reveal the PON's pX sequence and composition. The PONs that remain bound to the target oligonucleotide at the highest buffer temperature will most likely have the strongest binding affinity to the target, and are thus selected for further study and screening.

Once the primary construction of the PON is determined against a specific target, another round of construction and screening of PON sub-libraries is launched to determined the best secondary composition of the PON for achieving optimum biological activities *in vivo*. PON secondary libraries are sub-grouped according to the properties of the connecting AAs. Cationic PONs contain mostly lysine, arginine, etc as connecting AAs while anionic PONs generally consist of aspartic acid, glutamic acid, and their derivatives. PONs with serine, threonine, cystine, histidine, tyrosine, etc as major components tend to be more hydrophilic while those incorporating mostly valine, phenylalanine, tryptophan, proline, and 2,2-disubstituted amino acids as connecting AAs are more likely to be lipophilic. Each of those PON sub-libraries are screened for their ability to penetrate cell membranes in order to establish certain relationships between membrane penetration and the PON's chemical and physical properties such as lipophilicity and electronic charge. Based on the knowledge obtained from the experiments, new sub-libraries with mixed functionalities are constructed and screened until a proper combination of functionalities and sequences of connecting AAs is found best for the resulting PON to penetrate cell membranes and to bind to target sequences.

Screening of secondary PON libraries are generally performed in cell-lines carrying the gene of target sequence. Libraries are first labeled with ratio active amino acids at the N terminal of each PON molecule, and then released from resins. The soluble free PON mixtures are incubated with target cells at an appropriate temperature for a varying lengths of time and then aliquots of cells are taken at specific time, laid on the surface of 500 µl of pre-chilled

silicone oil, and centrifuged for 30 seconds in a Eppendorf centrifuge at ambient temperature. The bottom of the tube, which contains the cell pellet, is removed using dog toenail clippers, briefly inverted on absorbent paper to drain, and then transferred to a scintillation vial and counted for determining the apparent cell uptake of the PONs in the specific sub-library. After comparing apparent cell uptake of various PON sub-libraries, those having the most promising cell uptake are selected for further investigation.

Cells selected from above experiments are treated with detergents or physical forces such as-sonication and pressure to break the membranes. The total nucleic acids including DNAs and RNAs are isolated and digested with endonucleases. The antisense PONs that successfully penetrated the cell's membranes, reached to the active site, and bound specifically to the target would have formed doublex or triplex of PON-nucleic acid hybrids. These hybrids are resistant to nuclease digestion's and will remain as doublex and triplex of the same length as the starting PONs after the nuclease treatment. The PON-nucleic acid hybrids are then separated by electrophoresis on agarose gels. Since all the PONs, and thus their hybrids have the same nucleobase sequence and pX, the difference of their migration distance on the gel are directly determined by the type and sequences of connecting AAs in the binding PON. The strongest bands on the gel are cut. The hybrids are washed off from the gel and analyzed by Mass Spectrometry. Combining the information from mass spectrum, gel electrophoresis, and batch record of sub-library construction, the complete amino acid sequences of the selected antisense PONs can be determined.

EXAMPLES

The following examples are intended to illustrate, not to limit, the invention.

Example 1.

Synthesis of 3'-azido -2',3'-dideoxythymidine-5'-carboxylic acid:

Commercially available 3'-azido-2',3'-dideoxythymidine (AZT) (100 g, 0.38 mole) was dissolved in 1.5 L of 1.5 M aqueous postassium hydroxide (126 g KOH) solution with 4 equivalent of potassium persulfate ($K_2S_2O_8$, 370 g, 1.5 mole). With vigorous stirring, 0.1 equivalent of ruthenium trichloride ($RuCl_3$, 8.0 g, 0.038 mole) was added, and the solution temperature raised to 75 °C. The mixture was further stirred at room temperature for 16 hours, adjusted to pH=7 and concentrated under vacuum to near dryness. The residual solid was reslurried in fresh methanol (3 x 500 mL) and filtered. The methanol extracts were combined, de-colored with celite or silica gel and concentrated to about 400 mL to allow for crystallization. Crystals were filtered and dried to obtain 74.6 g (74%) of a light yellow solid: mp $\gtrsim$ 300 oC, ^{1}HNMR (DMSO-d6): δ 9.05 (s, 1H), 6.10 (1H, m), 4.42 (1H, m), 4.11 (1H,

m), 2.14 (2H, m), 1.75 (3H, s). ^{13}CNMR (DMSO-d_6): δ 175.2, 164.3, 150.8, 140.2, 106.9, 65.5, 65.2, 64.8, 36.9, and 13.0.

Example 2.

Synthesis of 3'-azido -2',3'-dideoxythymidine-5'-carboxylic acid ethyl ester:

The starting material 3'-azido -2',3'-dideoxythymidine-5'-carboxylic acid (10 g, 35.6 mmole) was dissolved in 400 mL of methanol. The solution was cooled to 0 oC in an ice bath, and 3.1 mL of thionyl chloride (5.1 g, 42.7 mmole) was added dropwise while maintaining the solution temperature below 5 oC. The solution was then stirred at room temperature for 10 hours, and concentrated to remove methanol. The residue was slurred in water (200 mL) and extracted with chloroform (3 x 100 mL). The chloroform extracts were combined, dried over anhydrous sodium sulfate and concentrated to dryness to obtain 8.0 g (76%) of a crystalline solid: mp = 125 - 128 °C, ^{1}HNMR ($CDCl_3$): δ 8.24 (1H. s), 6.43 (1H, m), 4.25 (4H, m), 2.23 (2H, m), 1.82 (3H, s), and 1.42 (3H, s). ^{13}CNMR ($CDCl_3$): δ 165.4, 161.6, 151.2, 141.2, 104.8, 71.4, 66.1, 64.9, 64.0, 37.5, 11.9 and 10.2.

Example 3.

Synthesis of 3'-amino -2',3'-dideoxythymidine-5'-carboxylic acid ethyl ester:

In a 500 mL bar bottle, the 3'-azido -2',3'-dideoxythymidine-5'-carboxylic acid ethyl ester (4.0 g, 13.6 mmole) was dissolved in 200 mL of methanol, and 200 mg of palladium/charcoal catalyst (10% Pd/C, dry) suspended in 2 mL of water was added. The mixture was shaken at room temperature under 40 psi of hydrogen for 4 hours and then filtered through celite. The cake was washed repeatedly with methanol (4 x 100 mL) and the filtrate was concentrated to dryness to obtain 3.1 g (85%) of a white solid: mp 148-150 oC, ^{1}HNMR (DMSO-d_6): δ 8.75 (1H. s), 6.33, 5.24, (2H, brs), 4.16 (1H, m), 4.01 (4H, m), 2.13 (2H, m), 1.83 (3H, s) and 1.51 (3H, s). 10.2.

Example 4.

Synthesis of 3'-amino -2',3'-dideoxythymidine-5'-carboxylic acid:

The starting material 3'-amino -2',3'-dideoxythymidine-5'-carboxylic acid ethyl ester (3.4 g, 12.6 mmole) was dissolved in 100 mL of ethanol (95%), and a solution of 1M sodium ethoxide (12.6 mL, 12.6 mmole) was added. The solution was stirred at room temperature overnight, neutralized with 1M hydrochloric acid (4.0 mL) to pH=7, and then concentrated to 40 mL. The mixture was cooled to 5 oC in an ice bath, filtered to remove NaCl precipitate, and concentrated to dryness to obtain 2.5 g (78%) of a while crystalline solid: mp > 300 oC,

^{1}HNMR (DMSO-d_6): δ 8.85 (1H. s), 6.33, (1H, m), 5.24, (2H, brs), 4.26 (1H, m), 3.81 (1H, m), 2.13 (2H, m), and 1.72 (3H, s).

Example 5.

Synthesis of BOC protected 3'-amino -2',3'-dideoxythymidine-5'-carboxylic acid:

In a typical reaction, the 3'-amino -2',3'-dideoxythymidine-5'-carboxylic acid (3.1 g, 12 mmole) was dissolved in 100 mL of 10% aqueous sodium carbonate solution, and cooled to 0 °C in an ice bath. A solution of di-tert-butyldicarbonate ($(Boc)_2O$, 3.1 g, 14 mmole) in dioxane (50 mL) was added slowly in a period of 1 hour. The reaction mixture was further stirred at room temperature for 16 hours, diluted with water (100 mL), and extracted with diethyl ether to remove byproducts and impurities. The aqueous phase was acidified with concentrated hydrochloric acid to pH=2 to produce a precipitation. The solid was collected by filtration and then recrystallized in nitromethane to obtain 3.4 g (68%) of white crystals: mp > 300 oC, ^{1}HNMR (DMSO-d_6): δ 11.24, (H, brs), 8.65 (1H. s), 6.13, (1H, m), 4.06 (1H, m), 3.96 (1H, m), 2.03 (2H, m), 1.72 (3H, s) and .

Example 6.

Synthesis of α-amino-γ-bromobutyric aicd hydrogenbromide:

Racemic homoserine (10 g, 84 mmole) was stirred in a 500 mL pressure tube with 30% (HBr/HOAc, w/w) hydrogen bromide in acetic acid (100 mL, 50 mmole HBr). The tube was sealed and then heated slowly in a water bath to 78 oC with good stirring. After holding at the temperature for 5 hours, the mixture was cooled down to room temperature, mixed with diethyl ether (100 mL), and filtered. The cake was washed with ether (3 x 50 mL), and dried in air to obtain 20 g (92%) of a white solid: mp = 200 °C, ^{1}HNMR (CD_3OD): δ 4.08 (1H, t, J = 7.8 Hz), 3.57 (2H. m), 2.46, (1H, m), and 2.28 (1H, m). The product can also be obtained from homoserine γ-lactone by following the same procedure.

Example 7.

Synthesis of ethyl α-amino-γ-bromobutyrate hydrogenchloride:

In a typical reaction, α-amino-γ-bromobutyric acid hydrogenbromide (7.4 g, 28.1 mmole) was dissolved in 100 mL of absolute ethanol. The solution was cooled to 5 oC in an ice bath, and slowly bubbled with gaseous hydrogen chloride (HCl) continuously for 8 hours. After further stirring under HCl at room temperature overnight and checking by ^{1}HNM for completion of the conversion, the mixture was concentrated to dryness. The residue was triturated with diethyl ether to produce a precipitate which was filtered, and dried to obtain 7.0

g (100%) of a white solid: mp = 200 °C, ^{1}HNMR (DMSO-d_5): δ 8.56 (3H, br.s, NH_3^-), 4.21 (2H. q, J = 6.15 Hz), 4.10, (1H,br.s), 3.62 (2H. m), 2.34 (2H, m) and 1.22 (3H, t, J = 7.0).

Example 8.

Synthesis of ethyl α-t-butoxycarbonylamino-γ-bromobutyrate:

Anhydrous sodium carbonate (6.4 g, 60 mmole) was added to 200 mL of dioxane solution containing di-ter-butyldicarbonate (6.2 g, 28.5 mmole) and ethyl α-amino-γ-bromobutyrate hydrogenchloride (7.4 g, 30 mmole). The suspension was stirred at room temperature overnight and then filtered. The solid was washed with dioxane and the filtrate was concentrated to dryness. The oily residue was reslurried in 200 mL of water containing 5% citric acid, and extracted with chloroform (3 x 100 mL). The chloroform extracts were combined, dried over anhydrous sodium sulfate, and concentrated to dryness to obtain an oil which was further dried in vacuum oven to yield a white crystalline solid: mp = 140 °C, ^{1}HNMR (DMSO-d_6): δ 5.12 (1H, br.s, NH), 4.38 (1H,br.s), 4.18 (2H. q, J = 7.15 Hz), 3.40 (2H, m), 2.34 (2H, t, J = 6.85), 2.39 (1H, m), 2.18 (1H. m), 1.41 (9H, s, t-Bu) and 1.25 (3H, t, J = 7.15).

Example 9.

Synthesis of ethyl α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate:

Ethyl α-t-butoxycarbonylamino-γ-bromobutyrate (4.5 g, 14.5 mmole), thymine (7.3 g, 58 mmole), and anhydrous potassium carbonate (k2CO3) (2.0 g, 14.5 mmole) was added to 100 mL of dry dimethylsufoxide (DMSO) pre-heated to 100 oC. The mixture was stirred at 100 oC for 4 hours and then concentrated to dryness under high vacuum. The residue was reslurried in 200 mL of chloroform, stirred at room temperature overnight, and filtered and washed repeatedly with chloroform (3 x 100 mL). The chloroform extracts were combined, dried over anhydrous sodium sulfate, and concentrated to dryness to obtain an oil. Both TLC and 1HNMR indicate that the crude oil contains a mixture of N1 and N3 alkylated thymine in a ratio of 60 to 40. This mixture was further purified by silica gel (Merck grade 10181, 35-70 mesh) column (2.5 x 50 cm glass column) chromatography. The column was first eluted with 1500 mL of hexane/ethyl acetate/ triethyl amine (10 : 20 : 1.5), then with hexane/ethyl acetate/ triethyl amine (10 : 10 : 1). An average volume of 250 mL for each fractions was collected. Fractions 9-12 were combined and concentrated to dryness to obtain a white solid which was recrystallized in methanol/water (2:1) to yield 1.5 g (30%) of pure ethyl α-t-

butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate: mp = 140 °C, ^{1}HNMR ($CDCl_3$): δ 8.69 (1H, br.s, NH on thymine), 7.07 (1H,br.s, thymine-CH_3), 5.32 (1H. d, J = 8.1 Hz, NHBoc), 4.29 (1H, m), 4.16 (2H, t, J = 6.85), 3.92 (1H, m), 3.66 (1H, m), 2.21 (1H, m), 2.00 (1H, m), 1.89 (3H, s), 1.43 (9H, s, t-Bu) and 1.25 (3H, t, J = 7.05).

Example 10.

Enzymatic resolution of racemic ethyl α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate

Synthesis of optically pure α(S)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acid:

In a typical reaction, the starting ester (1.5 g, 4.2 mmole) was dissolved in 50 mL of a mixture of water and acetonitrile (80 : 20) with 1,500 units of commercial protease papain (Sigma, 50 mg). The mixture was stirred at room temperature for 4 hours and the progress of the reaction was monitored by chiral HPLC. When about 50% of one enantiomer ester is converted to the corresponding acid, the mixture was extracted repeatedly with diethyl ether (4 x 30 mL). The ether extracts were combined, dried over anhydrous sodium sulfate, and concentrated to dryness to obtain 0.8 g (53%) of optically active ethyl α(R)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate. The aqueous phase was dried by liopholization. The residue was reslurried in absolute ethanol (50 mL), heated to 60 oC, then cooled to 5 oC and filtered. The filtrate was concentrated to small volume (10 mL) for crystallization. The solid was filtered, and dried to obtain 0.6 g (43%) of the optically pure α(S)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acid: ee = 98.7% (HPLC), mp = 240 °C, ^{1}HNMR (D_2O): δ 7.48 (1H,br.s, thymine-CH_3), 3.72-3.92 (3H, m), 2.15 (1H, m), 1.98 (1H, m), 1.82 (3H, s), and 1.36 (9H, s, t-Bu). MS (m/z): 327 (M+), 312 (M-CH3), 299 (M-CO), 271 (M-2CO), 255, 227, 152, 126, 86, 58.

For the analogous enzymatic resolution of the reverse reaction carried out in organic solvents, racemic α-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid (0.4 g, 1.1 mmole) was dissolved in 10 mL of 1 M citrate-phosphate buffer, pH = 4.2 with papain (100 mg). A mixture of ethanol (4 mL) and hexane (mL) was added to a final ratio of ethanol/hexane/buffer (1/1/3). The mixture was stirred

vigorously at room temperature for 24 hours and the progress to the reaction was monitored by HPLC. After about 50% acid to ester conversion, the reaction mixture was filtered to remove insoluble proteins, diluted with water, and extracted repeatedly with hexane. The hexane extracts was combined, washed with water, dried, and concentrated to dryness to obtain the optically pure active α-(S)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyrate. The aqueous layer is concentrated under vacuum to near dryness followed by crystallization to give optically pure α-(R)-amino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)-N-*t*-butoxycarbonylbutyric acid.

Synthesis of optically active α(R)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acid:

The optically active ethyl α(R)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate (0.8 g, 2.2 mmole) obtained from enzymatic resolution was dissolved in 20 mL of methanol, and 3.0 mL of 1M sodium hydroxide solution was added. The mixture was stirred at room temperature overnight, adjusted to pH = 7 with 1M hydrochloric acid, and then concentrated to dryness. The residue was reslurried in absolute ethanol (20 mL), heated to 60 oC, then cooled to 5 oC and filtered. The filtrate was concentrated to small volume (10 mL) for crystallization. The solid was filtered, and dried to obtain 0.66 g (90%) of the optically pure α(S)-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acid: ee = 90.3%, mp = 240 °C, ^{1}HNMR (D_2O): δ 7.48 (1H,br.s, thymine-CH_3), 3.72-3.92 (3H, m), 2.15 (1H, m), 1.98 (1H, m), 1.82 (3H, s), and 1.36 (9H, s, t-Bu). MS (m/z): 327 (M+), 312 (M-CH3), 299 (M-CO), 271 (M-2CO), 255, 227, 152, 126, 86, 58.

Chiral HPLC separation method:

(S)- and (R)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acids were separated on a CHIROBIOTIC T™ chiral HPLC column (250 x 4.6 mm, Astec) eluted with a solvent system of methanol and 1% triethylamine acetate buffer (pH = 4.0) in a ratio of 20 to 80. The flow rate was adjusted to 1.0 mL/min. with a back pressure of 2400 psi. The chromatogram was monitored by continuous wave length UV spectrometer on a HP-1090 Liquid Chromatography. Samples of a concentration about 1 mg/mL in the eluting solvent were injected in 10 µl volumes. The above solvent system separates a variety of enantiomers

of both N-protected and free a-amino butyric acid nucleosides, but not their ethyl esters. Of a typical run on this system, the retention times for (S)- and (R)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric, and for racemic ethyl α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyrate are 8.1, 10.6, and 14.8 minutes respectively. The detection limit for enantiomeric impurities of each enantiomers was below 0.5%.

Example 11.

Construction of Peptidoligonucleotides (PONs) through Standard Solid Phase Peptide Synthesis

Synthesis of a PON with 10 repeating (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric (pX) and 10 glysine (AA) connected through amide linkages as a 20-amino acid-residue peptide $(pX\text{-}AA)_{10}$.

The protocols chosen for construction of the PONs in this invention was based on standard Boc chemistry of solid phase peptide synthesis. In this specific example, 0.1 mM of Boc-Gly-MBHA (p-methylbenzhydrylamine) resin was used as starting material on an automated peptide synthesizer. After deprotection of the Boc-group with trifluoroacetic acid (TFA), the resin was coupled with optically pure (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric in the presence of 2-(1H-benzotriozole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU). After deprotection and washing, Boc-glycine is coupled to the resin using the same coupling reagents followed by deprotection, washing, and coupling of (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric again. The cycle was repeated until the 10th (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric was attached to the growing peptide chain. The resin was then deprotected and coupled with L-lysine as the N-terminal residue to increase the water solubility of the resulting PON molecule. The peptide was then cleaved from the MBHA resin and purified by preparative HPLC to obtain 11 mg of the tittle PON as a off-white powder, with lysine at the N terminal and glycine amide at the C-terminal ($H_2N\text{-}(gly\text{-}pT)_{10}\text{-}lys$, p = 2-aminobutyric acid). This peptido oligonucleotide is >93% pure by RP-HPLC:

Instrument: Beckman System Gold

	Shimadzu CR4A Intergrator
Column:	Vydac C18 218TP104
Solvent A:	0.1% (W/V) TFA/H_2O
Solvent B:	0.1% (W/V) TFA/CH_3CN
Gradient:	5-60% B in 27.5 minutes
Flow rate:	1.0 mL/min.
Wavelength:	215 nm
Product Rt:	16.14 min.

Mass Analysis (Ion Spray): Molecular Weight = 2806.1, (M + H)+ = 2805.8, and (M + Na)+ = 2825.7.

Example 12.

Synthesis of resin-bind PON combinatorial libraries based on one-bead-one-peptide strategy.

A 10-nucleobase library was constructed with (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric for all pXs, and different combinations of glycine, alanine, phenylalanine, lysine, and aspartic acid as connecting AAs. Standard solid phase peptide synthetic procedures are followed as in example 11. Boc-glycine MBHA resins were deprotected, and coupled with (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric (pX). After deprotection of the Boc group on the pX, the resins were equally divided into four proportions and each was coupled respectively with properly protected Boc- alanine, phenylalanine, lysine(Cl-Z), and aspartic acid (OBzl). These resins were then combined, thoroughly mixed, deprotected, and then coupled with (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric acid (pX) again. This cycle of coupling the pX, dividing the resins, coupling the AAs, combining the resins, and coupling the pX again is repeated until the 10th pX is incorporated. The resins containing a theoretical number of 262,144 different individual PONs with the same $(pT)_{10}$, nucleotide sequence were washed, dried, and re-suspended in 10 mM phosphate buffer (100 mM NaCl, pH = 7.2). Deoxythymidine 10-mer labeled with fluorescent reagent 4-acetamido-4'-isothiocyanatostibene-2,2'-disulfonic acid at the 5' end was added to the mixture. and the suspension was heated to 90 °C under vigorous stirring. After slowly cooling to 4 °C, The mixture is re-heated to 80 °C and filtered. The resins are washed repeatedly with hot (80 °C) buffer (10 mM phosphate, 100 mM NaCl, pH = 7.2), and plated under UV light for

visualization. Resin beads emitting strong fluorescence are picked and decoded to reveal the peptide sequences.

Example 13.

Synthesis of soluble PON combinatorial libraries for in vivo screening.

The same 10-nucleobase library was constructed with (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric for all pXs, and different combinations of glycine, alanine, phenylalanine, lysine, and aspartic acid as connecting AAs. Same standard solid phase peptide synthetic procedures are followed as in example 11. However, instead of using the one-bead-one-peptide strategy, the above amino acids are coupled to the resins as mixtures in a predetermined ratio. Boc-glycine MBHA resins are swollen in DMF and dichloromethane (DCM), deprotected with TFA, neutralized with diisopropylethylamine in DCM, and then coupled with (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric using HBTU as the coupling reagent. After deprotection, neutralization and wash, the resins are coupled with a mixture of Boc-Ala, Boc-Phe, Boc-lys(Cl-Z), and Boc-Asp(OBzl) in a ratio of 1 : 1.15 : 1.56 : 1.23. The cycle of alternately coupling (S)-α-t-butoxycarbonylamino-γ-(1-(2,4-dihydroxy-5-methylpyrimidyl)butyric (pX) and the four Boc-amino acid mixtures is continued until the 10th pX is incorporated. The resins are deprotected, and coupled with ^{14}C-glycine to introduce radioactivity. The labeled PON mixtures are then released from the resins by HF cleavage, the HF was removed under nitrogen stream, and the products are dried under vacuum in a desiccator overnight. The free peptides are then re-dissolved in appropriate buffer for in vivo screening. The solution of the PON library is incubated with target cells at a proper temperature for a varying lengths of time and then aliquots of cells are taken at specific time, laid on the surface of 500 µl of pre-chilled silicone oil, and centrifuged for 30 seconds in a Eppendorf centrifuge at ambient temperature. The bottom of the tube, which contains the cell pellet, is removed using dog toenail clippers, briefly inverted on absorbent paper to drain, and then transferred to a scintillation vial and counted for determining the apparent cell uptake of the PONs in the specific sub-library. After comparing apparent cell uptake of various PON sub-libraries, those having the most promising cell uptake are treated with detergents or physical forces such as sonication and pressure to break the membranes. The total nucleic acids including DNAs and RNAs are isolated and digested with endonucleases. The PON-nucleic acid hybrids are separated by electrophoresis on agarose gels. The strongest bands on the gel are cut, washed off from the gel with strong ionic buffer, and analyzed by Mass Spectrometry. Combining the information from mass

spectrum, gel electrophoresis, and batch record of sub-library construction, the complete amino acid sequences of the selected antisense PONs are determined.

Example 14.

Recognition of deoxyadenasine 12-mer $(dA)_{12}$ by PON (lys-$(pT\text{-}gly)_{10}$-gly-NH_2).

An increase in UV absorbance is observed during the thermal denaturation as the ordered, native structure of a nucleic acid base-pair stacking is disrupted. Known as hypochromicity, the change in UV absorbance is a measure of base-pairing and base-stacking between two complementary strands. The resulting UV absorbance profile as a function of temperature is known as a melting curve with the midpoint of the curve defining the melting temperature, T_m, at which 50% of the double strand is dissociated into its two single strands. The measurement of UV absorbance melting curves provides qualitative and quantitative structural information about the nucleic acid bound to its complementary strand. The T_m is dependent upon the concentration of the oligonucloetide the properties to the solvent (buffer: pH, ionic strength, ect.).

Binding studies were carried out by hybridizing the PON described above to its complementary oligonucleotide "dA_{12}", followed by thermal denaturation and measurement of the UV absorbance as a function of temperature. Synthetic oligodeoxynucleotide (ODN) dT_{12} and its complementary oligodeoxynucleotide dA_{12} were used as reference nucleic acids. The samples were prepared in 50mM phoshate (Na_2HPO_4)(pH 7.4) and 140 mM NaCl buffer at 5 mM dA_{12}, 10 mM dT_{12} and 10 mM PON T_{10}.. Aliquots (0.5 mL) of A_{12} and PON T_{10} or A_{12} and dT_{12} were mixed in Eppondorf tubes and transferred to 1mL cuvettes. Samples were heated from 15 °C to 95 °C at a rate of 0.5 °C/min. The change in absorbance was measured over the heating period. Results of the melting point determination is shown in TABLE 1 To test the binding strength of the PON due to mismatched nucleotides within the target strand, synthetic oligodeoxynucleotides containing a single and double T were made and binding studies were carried out with the PON described above. The results are summarized in TABLE 1.

These data confirm that the PON binds to its complementary nucleic acid strand, and that it exhibits a transition from an ordered structure to a disordered one on thermal denaturation.

The T_m of the PON:ODN was significantly higher than that of the control ODN dimer dA_{12} : dT_{12}. This indicates very strong interaction between the two strands. Even with single and double mismatch ODNs (dA_{11} : dT_1 , dA_{10} : dT_2) the T_m was still higher than the control ODN dimer. This adds support to the binding interaction of the PON to the ODN and that this interaction is due to base-pair interaction rather than a non-specific interaction.

TABLE 1:

Reference	Complementary Strand		Stiochio-metry	T_m (°C)
dNu	dNu	pNu		
dA_{12}	$2dT_{12}$	-	1:2	25.2
dA_{12}	-	$2pT_{10}$	1:2	74.3
dA_{11} dT_1	-	$2pT_{10}$	1:2	58.7
dA_{10} dT_2	-	$2pT_{10}$	1:2	49.2

dNu = deoxyoligonucleotide; pNu = Peptido-Oligonucleotide; pT_{10} = PON T_{10}

CONCLUSIONS

The most innovative and unique feature in this invention of peptido oligonucleotide (PON) and its libraries is the introduction of an element of plurality into the oligonucleotides by alternately connecting an amino acid and a nucleoside through peptide synthesis. The resulting peptide oligonucleotides are not only highly analogous to nucleic acids in terms of recognizing and base-pairing with complementary sequences, and possessing the peptide backbone that resist nuclease degradation, but also have the flexibility of carrying various combinations of functionalities within the molecule without changing the nucleotide sequence or increasing the length of the peptide chain. By simply selecting different amino acids in each coupling step for the connecting AAs during peptide synthesis, a great variety of PONs can be generated against a specific target without adding extra steps. This unique feature renders the peptide oligonucleotides ideal candidates for combinatorial library construction. Combining the feature with a powerful screening method also disclosed in this invention, desirable antisense oligonucleotides that (1) can be synthesized easily and in bulk; (2) are stable *in vivo*; (3) can

effectively enter the target cell; (4) can be retained by the target cell; and (5) can bind strongly and specifically to cellular targets; could be selected in a much shorter time frame.

The potential utility of the peptide oligonucleotides (PONs) of this invention is far reaching. The ability of antisense oligonucleotides to recognize and bind to specific sequences in a DNA or RNA molecule is the foundation for their wide spread applications. Extensive research and development work has been carried out in using antisense oligonucleotides for treatment or diagnoses of gene related diseases, especially cancer, AIDS, and other genetic disorders. Besides the utilities applied by other antisense oligonucleotides, the PONs of this invention can be further explored in other applications outside traditional antisense arena. These PONs are designed in a way that allows to easily incorporate various functional groups in the backbone of the oligomers. While these oligomers recognize and bind to specific gene sequences, the functional groups on their backbones can serve as catalytic arms that reach across to the complementary strain and perform certain chemical reactions. This combination of sequence specificity with catalytic activity in one synthetic oligonucleotide provides a perfect tool to design artificial enzymes for gene surgery. DNAs and RNAs could potentially be cleaved, ligated, alkylated, oxidized, reduced, halogenated, etc. on any base at any desired sequence catalyzed by these specifically designed *catalytic PONs*. These synthetic *catalytic PONs* could also be used as probes for mechanistic studies of a vast variety of DNA or RNA processing enzymes such as nucleases, ligases, RNAases, and polymerases.

One example of the application of these *catalytic PONs* is the design and synthesis of artificial sequence-specific endonucleases. Restriction endonucleases are extremely important tools in molecular biology and biochemistry. They are widely used in gene isolation, DNA sequencing, and recombinant DNA technology. However, these enzymes recognize relatively small size DNA sequences, (usually 4-6 bp) and thus generate too many fragments from a large DNA substrate. On the other hand, there is only a limited number of restriction enzymes currently available and many of them have overlapping specificity's. There remain numerous sequences for which no restriction enzymes are available. The synthetic *catalytic PONs* would clearly have the potential capability of cleaving DNA and RNA with high specificity at any desired site and would thus provide a valuable tool to molecular biologists.

The above descriptions is only illustrative and not restrictive. Those skilled in the art will appreciate that numerous changes and modifications can be made without departing from the spirit of the invention. It is therefore intended that the appended claims cover all such equivalent variations as fall within the true spirit and scope of the invention.

What is claimed is:

Other Publications

Agrawal, Sudhir. "Antisense oligonucleotides: towards clinical trials. Trends Biotechnol., 1996, 14, 376-387.

Egli, Martin. "Structural aspects of nucleic acid analogs and antisense oligonucleotides. Angew. Chem., Int. Ed. Engl., 1996, 35, 1895-1909.

Altmann, K.-H.; Fabbro, D.; Dean, N. M.; Geiger, T.; Monia, B. P.; Mueller, M.; Nicklin, P. "Second-generation antisense oligonucleotides: structure-activity relationships and the design of improved signal-transduction inhibitors". Biochem. Soc. Trans., 1996, 24, 630.

Dean, N. M.; McKay, R.; Miraglia, L.; Geigert, T.; Mullert, M.; Fabbro, D.; Bennett, C. F., "Antisense oligonucleotides as inhibitors of signal transduction: development from research tools to therapeutic agents.", Biochem. Soc. Trans., 1996, 24, 623-629.

Zelphati, Olivier; Szoka, Francis C., Jr., "Liposomes as a carrier for intracellular delivery of antisense oligonucleotides: a real or magic bullet". J. Controlled Release., 1996, 41, 1,2, Fifth International Symposium on Delivery and Targeting of Pesticides, Proteins and Genes, 1995, 99-119.

Crooke, Stanley T. "Proof of mechanism of antisense drugs.", Nucleic Acid Drug Dev., 1996, 6, 145-147.

Baba, Masanori. "Development of antisense drugs. Antisense oligonucleotides as anti-AIDS agents." Organ Biol., 1995, 2, 13.

Plenat,-Francois. "Animal models of antisense oligonucleotides: Lessons for use in humans." Mol. Med. Today, 1996, 2, 250-257.

Ma, L.; Calvo, F. "Recent status of the antisense oligonucleotide approaches in oncology." Fundam. Clin. Pharmacol., 1996, 10, 97-115

Hanvey, Jeffery C.; Babiss, Lee E.,"Peptide nucleic acids as antisense therapeutic agents", Delivery Strategies Antisense Oligonucleotide Ther., 1995, 151-60.

De Mesmaeker, Alain; Altmann, Karl-Heinz; Waldner, Adrian; Wendeborn, Sebastian, "Backbone modifications in oligonucleotides and peptide nucleic acid systems, Curr. Opin. Struct. Biol., 1995, 5, 343-55.

Noble, Stewart A.; Bonham, Michele A.; Bisi, John E.; Bruckenstein, David A.; Brown, Pamela H.; Brown, Stephen C.; Cadilla, Rodolfo; Gaul, Micheal D.; Hanvey, Jeffrey C.; et al., "Impact of biophysical parameters on the biological assessment of peptide nucleic acids, antisense inhibitors of gene expression. Drug Dev. Res. 1995, 34, 184-95.

Takao, Toshifumi; Fukuda, Hiroyuki; Coull, James; Shimonishi, Yasutsugu., "Positive-ion fast-atom bombardment tandem mass spectrometry of peptide nucleic acids. Rapid Commun. Mass Spectrom. 1994, 8, 925-8.

Almarsson, Orn; Bruice, Thomas C. "Peptide nucleic acid (PNA) conformation and polymorphism in PNA-DNA and PNA-RNA hybrids." Proc. Natl. Acad. Sci. U. S. A., 1993, 90, 9542-6.

Rose, Donald J. "Characterization of antisense binding properties of peptide nucleic acids by capillary gel electrophoresis. Anal. Chem. 1993, 65, 3545-9.

Lam, Kit S.; Lebl, Michal. "Peptide library screening based on the one-bead one-peptide concept." Biol. Chem., Proc. Chin. Pept. Symp., 3rd, 1995, 273-277.

Pinilla, C.; Appel, J. R.; Blondelle, S. E.; Dooley, C. T.; Dorner, B.; Eichler, J.; Ostresh, J. M.; Houghten. "Synthetic combinatorial libraries screened in solution." Chem., Struct. Biol., Proc. Am. Pept. Symp., 14th, 1996, 884-886.

Dooley, C. T.; Bower, A. N.; Houghten, R. A. "Identification of Mu-selective tetrapeptides using a positional scanning combinatorial library containing L-, D- and unnatural amino acids." Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 14th, 1996, 623-624.

Sebestyen, F.; Kindla, K.; Rapp, W.; Campian, E.; Furka, A. "Synthesis of colored peptide libraries." Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 14th, 1996, 319-320.

Blondelle, S. E.; Forood, B.; Houghten, R. A.; Perez-Paya, E. "Design of conformationally defined combinatorial libraries based upon "protein-like" structural motifs." Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 14th, 1996, 293-294.

Lam, Kit S.; Lebl, Michal. "Synthesis and screening of a "one -bead-one-compound" combinatorial peptide library.", Methods Mol. Cell. Biol., 1996, 6, 46-56.

Doerner, Barbara; Ostresh, John M.; Husar, Gregory M.; Houghten, Richard A. "Generation of peralkylated peptidomimetic combinatorial libraries." Methods Mol. Cell. Biol.. 1996, 6, 17-22.

Nawrocki, Joseph P.; Wigger, Maria; Watson, Clifford H.; Hayes, Thomas W.; Senko, Michael W.; Benner, Steven A.; Eyler, John R. "Analysis of combinatorial libraries using electrospray Fourier transform ion cyclotron resonance mass spectrometry." Rapid Commun. Mass Spectrom., 1996, 10, 1860-1864.

Winger, Brian E.; Campana, Joseph E. "Characterization of combinatorial peptide libraries by electrospray ionization Fourier transform mass spectrometry." Rapid Commun. Mass Spectrom.","1996, 10, 1811-1813.

Houghten, R. A.; Dooley, C. T.; Ostresh, J. M. "The use of soluble polyamine combinatorial libraries for the identification of potent opioid receptor active compounds." Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 14th, 1996, 278-280.

Vandersteen, Anthony M.; Han, Hyunsoo; Janda, Kim D. "Liquid-phase combinatorial synthesis: in search of small-molecule enzyme mimics.", Mol. Diversity, 1996, 2, 89-96.

Salmon, Sydney E.; Liu-Steven, Rosa H.; Zhao, Yu; Lebl, Michal; Krchnak, Viktor; Wertman, Ken; Sepetov, Nikolai; Lam, Kit S. "High-volume cellular screening for anticancer agents with combinatorial chemical libraries: a new methodology." Mol. Diversity, 1996, 2, 57-63.

Houghten, Richard A.; Blondelle, Sylvie E.; Dooley, Colette T.; Doerner, Barbara; Eichler, Jutta; Ostresh, John M. "Libraries from libraries: generation and comparison of screening profiles." Mol. Diversity, 1996, 2, 41-45.

Sepetov, N. F.; Krchnak, V.; Stankova, M.; Lam, K. S.; Wade, S.; Lebl, M. "New approach in constructing and screening peptide libraries." Pept. 1994, Proc. Eur. Pept. Symp., 23rd, 1995, 467-468.

Lam, K. S.; Lebl, M.; Wu, J.; Wade, S.; Lou, Q.; Zhao, Z. G.; Chen, M. L.; Chang, M. P.; Smith, M.; Grimes, W. J. "Peptide library screening based on the one-bead one-peptide concept: An update." Pept. 1994, Proc. Eur. Pept. Symp., 23rd, 1995, 453-454.

Antonenko, V. V.; Kulikov, N. V.; Tuelty, D.; Sundaram, A.; Jacobs, J. W. "Combinatorial libraries by light-directed, spatially addressable parallel chemical synthesis: Fidelity of the method and application to drug discovery.", Pept. 1994, Proc. Eur. Pept. Symp., 23rd, 1995, 451-452.

Nielsen, J.; Brenner, S.; Janda, K. D. "Implementation of encoded combinatorial chemistry.", Pept. 1994, Proc. Eur. Pept. Symp., 23rd, 1995, 92-93.

Felder, E. R.; Heizmann, G.; Matthews, I.; Rink, H.; Spieser, E. "An encoded synthetic library suitable for binding assays on beads and subsequent operations with selectively

Blondelle, S. E.; Takahashi, E.; Houghten, R. A.; Perez-Paya, E. "Design of conformationally defined combinatorial libraries.". Pept. 1994, Proc. Eur. Pept. Symp., 23rd, 1995, 85-86.

Bi, Yingzhi; Schultz, Peter G. "Building blocks for peptide and carbamate libraries.", Bioorg. Med. Chem. Lett., 1996, 6, 2299-2300.

Gao, Bin; Esnouf, M. Peter. "Elucidation of the core residues of an epitope using membrane-based combinatorial peptide libraries.", J. Biol. Chem., 1996, 271, 24634-24638.

Geysen, H. Mario; Wagner, Craig D.; Bodnar, Wanda M.; Markworth, Christopher J.; Parke, Gillian J.; Schoenen, Frank J.; Wagner, David S.; Kinder, Daniel S. "Isotope or mass encoding of combinatorial libraries.", Chem. Biol., 1996, 3, 679-688.

Nielsen, Peter E. "Peptide nucleic acids: a new dimension to peptide libraries and aptamers.", Methods Enzymol., 1996, 267, Combinatorial Chemistry, 426-433.

Ni, Zhi-Jie; Maclean, Derek; Holmes, Christopher P.; Gallop, Mark A. " combinatorial chemistry: binary coding using chemically robust secondary amine tags.", Methods Enzymol., 1996, 267, Combinatorial Chemistry, 261-272.

Ivanetich, Kathryn M.; Santi, Daniel V. "Preparation of equimolar mixtures of peptides by adjustment of activated amino acid concentrations.", Enzymol., 1996, 267, Combinatorial Chemistry, 247-260.

Janda, Kim D.; Han, Hyunsoo. "Combinatorial chemistry: a liquid-phase approach.", Methods Enzymol., 1996, 267, Combinatorial Chemistry, 234-247.

Ostresh, John M.; Blondelle, Sylvie E.; Dorner, Barbara; Houghten, Richard A. "Generation and use of nonsupport-bound peptide and peptidomimetic combinatorial libraries.", Enzymol., 1996, 267, Combinatorial Chemistry, 220-234.

Chen, Charlie L.; Strop, Peter; Lebl, Michal; Lam, Kit S. "One bead-one compound combinatorial peptide library: different types of screening.", Enzymol., 1996, 267, Combinatorial Chemistry, 211-219.

Blondelle, Sylvie E.; Perez-Paya, Enrique; Houghten, Richard A. "Synthetic combinatorial libraries: novel discovery strategy for identification of antimicrobial agents.", Antimicrob. Agents Chemother., 1996, 40, 1067-1071.

Harada, Kazuo; Martin, Shelley S.; Frankel. Alan D. "Selection of RNA-binding peptides in vivo.", Nature (London), 1996, 380, 175-9.

Boutin; Jean A.; Hennig, Philippe; Lambert, Pierre-Herve; Bertin, Sophie; Petit, Laurence; Mahieu, Jean-Pierre; Serkiz, Bernard; Volland, Jean-Paul; Fauchere, Jean-Luc. "Combinatorial peptide libraries: robotic synthesis and analysis by nuclear magnetic resonance, mass spectrometry, tandem mass spectrometry, and high-performance capillary electrophoresis techniques.", Anal. Biochem., 1996, 234, 126-41.

1. A stereochemically defined composition of peptido oligonucleotides (PONs) in the form of S-$(pX\text{-}AA)_n$-Y which possess superior properties as antisense agents for potential treatment of gene related diseases.

Wherein:

S is a hydrogen or a linker or a modifying group or a peptide.

Y is a hydrogen or a modifying group or an amino acid or a peptide.

AA is one of any natural and unnatural amino acids excluding pX.

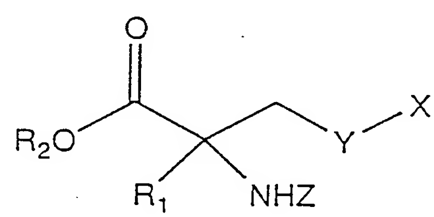
pX is an optically active amino acid nucleoside having the structure of $HOOCCHNH_2CH_2CH_2$-X, where X is any one of the nucleobase or their derivatives including thymine, cytosine, uracil, adenine, and guanine.

n = 1 or more.

2. A composition of claim 1, wherein said pX contains an (S) chrial center and said AA contains an (S), or an (R), or multiple (S) or (R), or defined multiple (S) and (R) chiral centers, or no chiral centers.

3. A composition of claim 1, wherein said pX contains an (R) chrial center and said AA contains an (S), or an (R), or multiple (S) or (R), or defined multiple (S) and (R) chiral centers, or no chiral centers.

4. A process for preparation of pure stereoisomers through enzymatic resolution of racemic mixtures of the composition



wherein:

X is a nucleoside base or its modified derivatives

R_1 is H; an alkyl or branched alkyl group; or a cyclic or heterocyclic ring system.

R_2 is H; an alkyl or branched alkyl group; or a cyclic or heterocyclic ring system.

Y is CH_2 or CH_2CH_2 or O or S.

Z is a protecting group including Fmoc, Boc, Cbz, Pht, etc.

5. The method as described in claim 4 comprising the steps of treating the racemic mixtures of the said composition where R2 is an alkyl group, with an hydrolytic enzyme to remove R2 enantioselectively, separating the hydrolyzed said composition from the unreacted starting material, and obtaining the optically active stereoisomers of the said composition.

6. The method as described in claim 4 comprising the steps of treating the racemic mixtures of the said composition where Z is an acyl group, with an hydrolytic enzyme to remove Z enantioselectively, separating the hydrolyzed said composition from the unreacted starting material, and obtaining the optically active stereoisomers of the said composition.

7. The method as described in claim 4 comprising the steps of treating the racemic mixtures of the said composition where R2 is an H, with an hydrolytic enzyme to make an ester derivative enantioselectively, separating the esterified said composition from the unreacted starting material, and obtaining the optically active stereoisomers of the said composition.

8. The method as described in claim 4 comprising the steps of treating the racemic mixtures of the said composition where Z is an H, with an hydrolytic enzyme to add an acyl group enantioselectively to the amine, separating the acylated said composition from the unreacted starting material, and obtaining the optically active stereoisomers of the said composition.

9. A method of synthesizing defined mixtures or combinatorial libraries of the composition S-(pX-AA)$_n$-Y

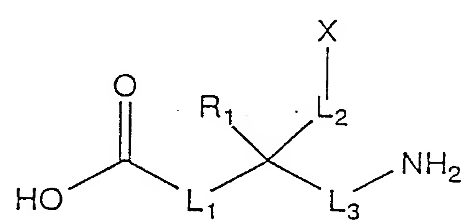
Wherein:

S is a hydrogen or a linker or an amino acid or a peptide.

Y is a hydrogen or a modifying group or a peptide

AA is any one of the natural and unnatural amino acid excluding pX.

pX is an optically active amino acid nucleoside having the structure of



wherein:

X is a nucleoside base or its modified derivatives

R1 is H; alkyl or substituted alkyl; alkenyl or substituted alkynyl; alkaryl or substituted alkaryl; aralkyl or substituted arakyl; alcyclic; cyclic and heterocyclic ring systems.

L1 is a bond or an atom or a group of atoms

L2 is a bond or an atom or a group of atoms

L3 is a bond or an atom or a group of atoms

L1, L2, L3, and R1 can be interconnected in one or more ring systems.

n = 4 or more.

10. The method as described in claim 9 wherein each mixture or library contains from several to millions of distinct, unique and different peptido oligonucleotides which all have the same nucleobase sequences, and recognize and bind to the same target nucleic acid.

11. The method as described in claim 9, comprising the steps of:

A. Coupling a pre-selected said pX to an activated substrate such as a resin or an amino acid-resin or a C-protected amino acid.

B. Removing the protecting group from said pX of the resulting substrate-pX chain.

C. Coupling a plurality of amino acids to the said substrate-pX to form a mixture or a group of mixtures of a new chain, substrate-pX-AA.

D. Removing the protecting group from said substrate-pX-AA, and repeating the process from step A to step D, until the desired length of the PON is reached, giving substrate-$(pX\text{-}AA)_n$.

12. The method as described in claim 9, comprising the steps of:

A. Coupling a plurality of said pX to a substrate or a group of substrates where p varies among different spacers but X is pre-selected according to the target sequence.

B. Removing the protecting group from said pX of the resulting substrate-pX mixtures.

C. Coupling an amino acid to the said substrate-pX to form a mixture or a group of mixtures of a new chain, substrate-pX-AA.

D. Removing the protecting group from said substrate-pX-AA, and repeating the process from step A to step D, until the desired length of the PON is reached, giving substrate-$(pX\text{-}AA)_n$.

13. The method as described in claim 7, comprising the steps of:

A. Coupling a plurality of said pX to a substrate or a group of substrates where p varies among different spacers but X is preselected according to the target sequence.

B. Removing the protecting group from said pX of the resulting substrate-pX mixtures.

C. Coupling a plurality of amino acids to the said substrate-pX to form further a mixture or a group of mixtures of a new chain, substrate-pX-AA.

D. Removing the protecting group from said substrate-pX-AA, and repeating the process from step A to step D, until the desired length of the PON is reached, giving substrate-$(pX\text{-}AA)_n$.

14. The method as described in claims 11 or 12 or 13, further comprising the steps of attaching to the N-terminal of the PON chain a ligand carrying certain functionalities such as radio or fluorescent labeling, metal chellating, chemical laminating, water hydrating, and nucleotide cleaving.

15. The method as described in claims 11 or 12 or 13, further comprising the steps of coupling more amino acids other than pX, or more peptides to the N-terminal of the PON chain in a combinatorial fashion.

16. The method as described in claim 7, further comprising the selection, isolation, and identification of a PON molecule having the desired chemical, physical, and biological properties.

17. The method as described in claim 16, comprising the steps of :

-Binding properly labeled target nucleic acid to a resin-bond PON library, Washing resin with continuously increased buffer temperature, select resins that remain bond to target at high temperature, and determine the composition of the PON on the resin.

18. The method as described in claim 16, comprising the steps of :

Treating a cell line carrying the target nucleic acid with a soluble PON library, isolating the total cell nucleic acids and digesting it with DNA and RNA nucleases, selecting double or triple stranded PON- nucleic acid complexes, and determining the compositions of the PON molecules.

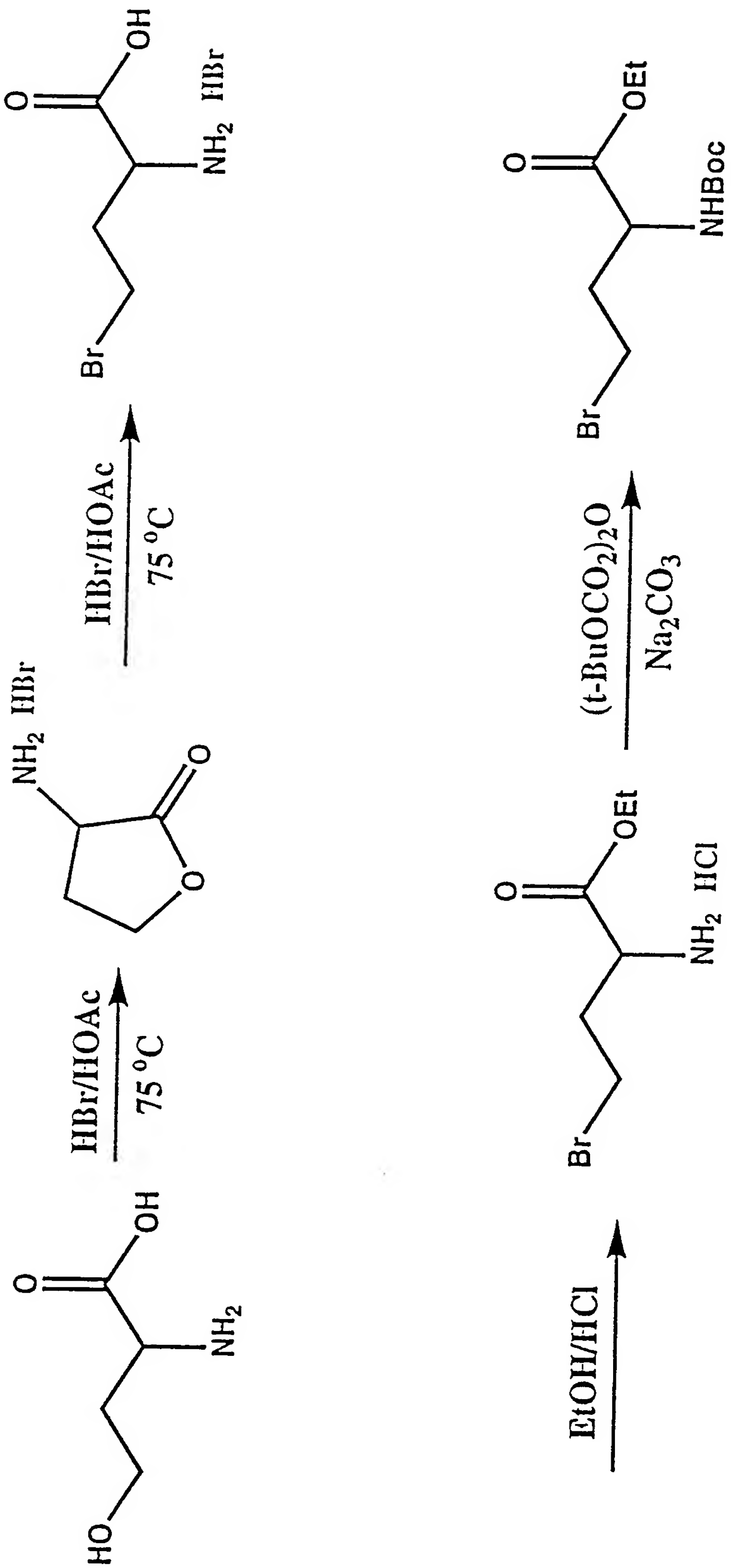


Fig. 1

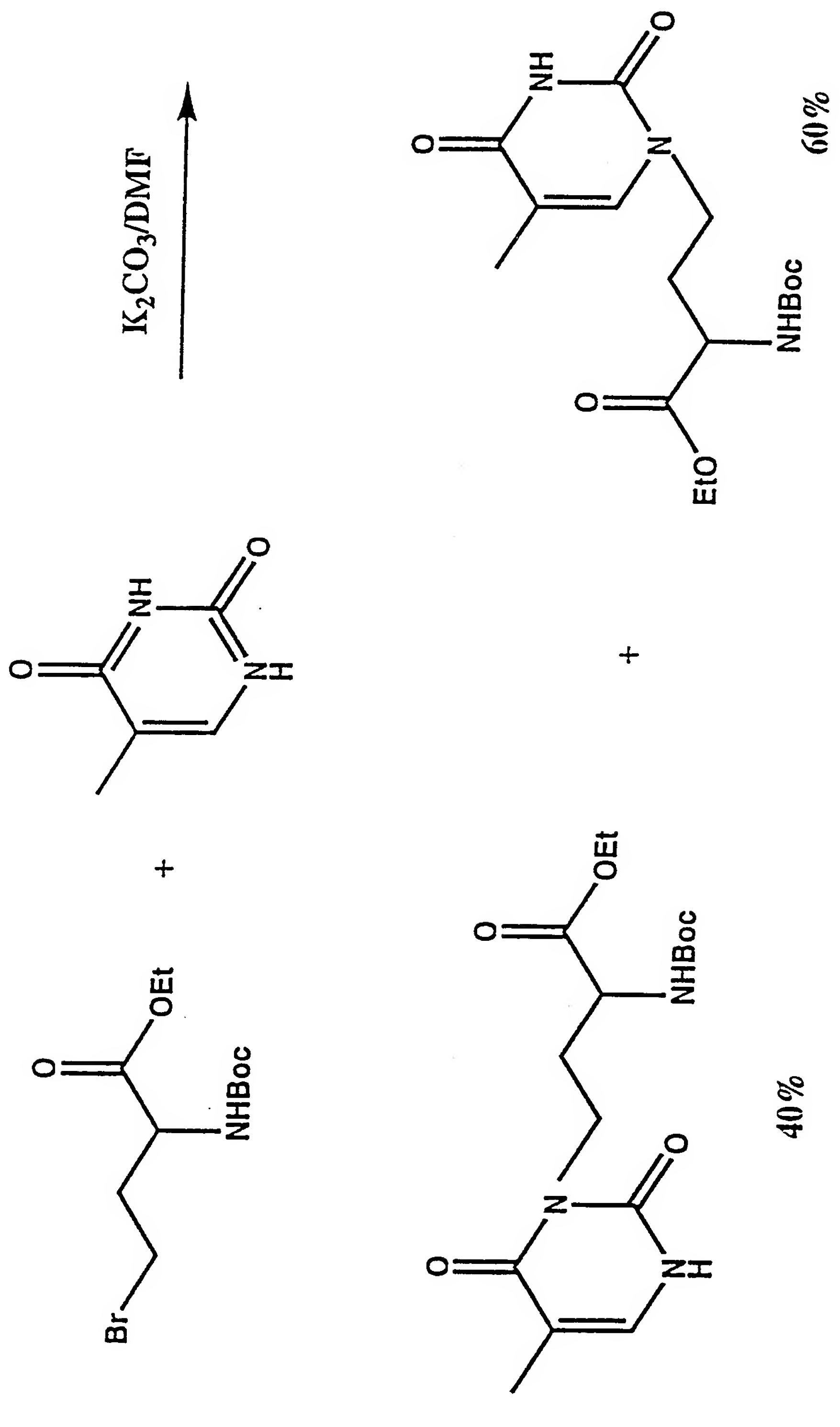

Br
NHBoc
O
OEt
+
O
NH
O
N
H
K_2CO_3/DMF
EtO
O
NHBoc
N
O
NH
O
60%
+
O
OEt
NHBoc
N
O
N
H
O
40%

Fig. 2

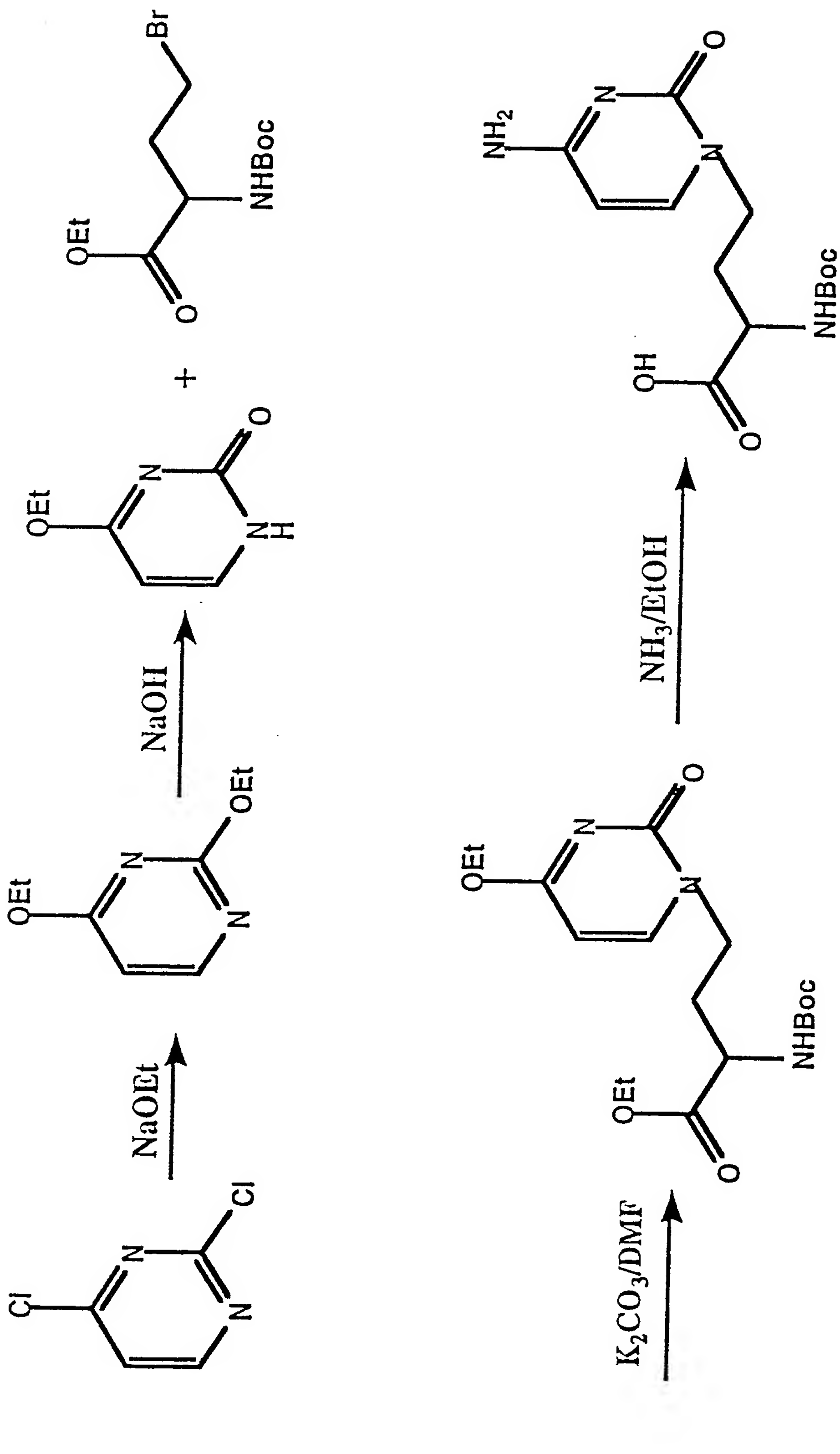


Fig. 3

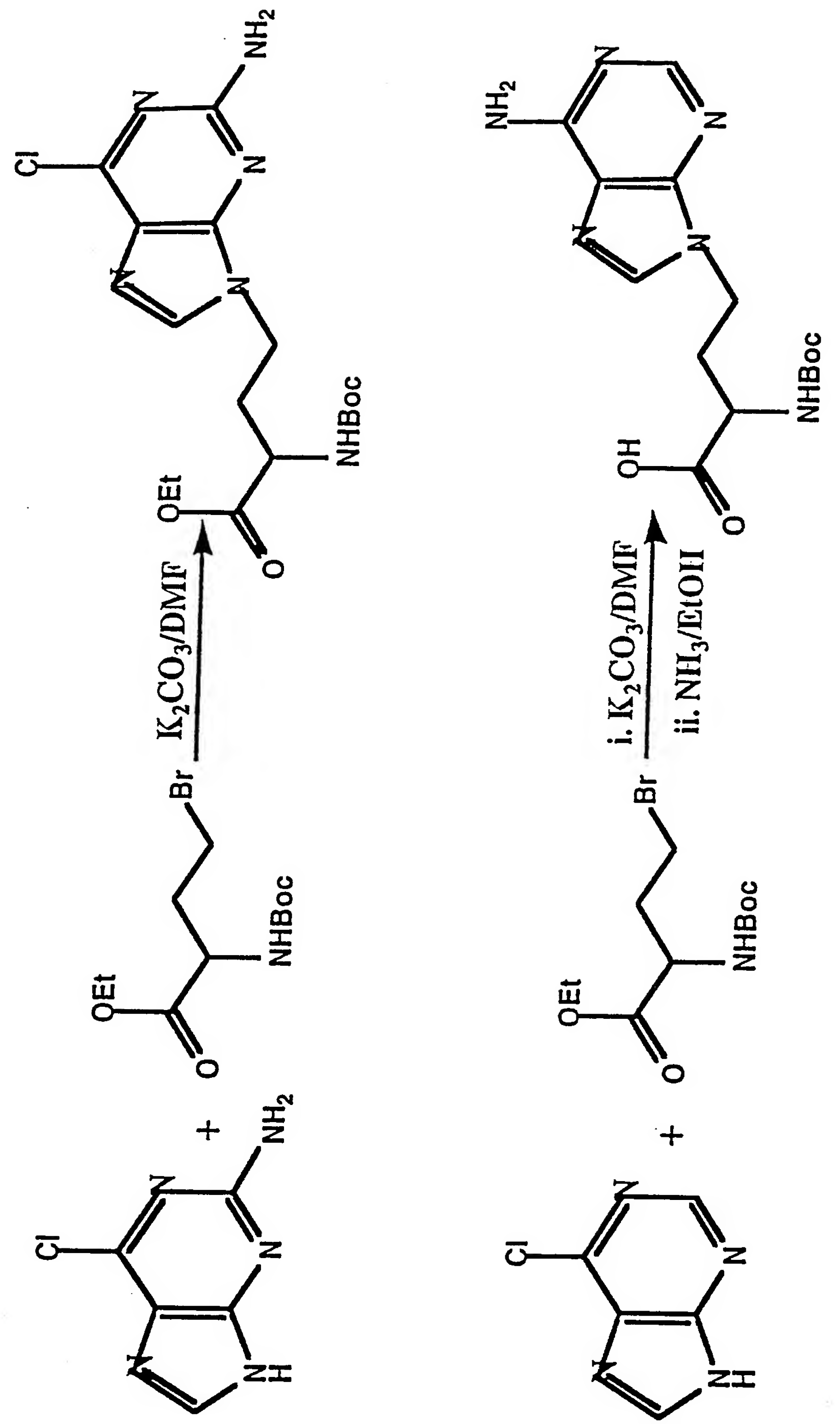


Fig. 4

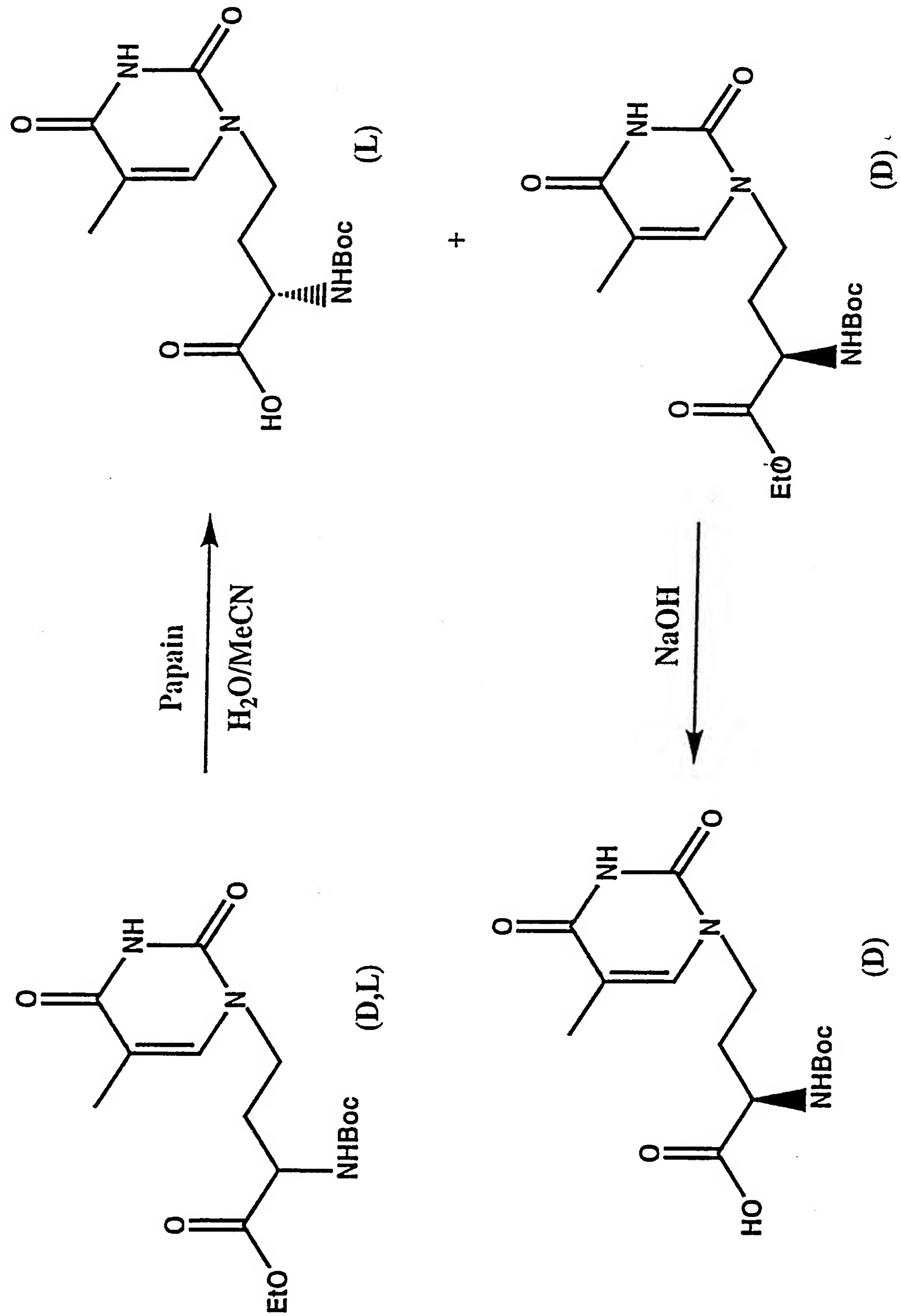


Fig. 5

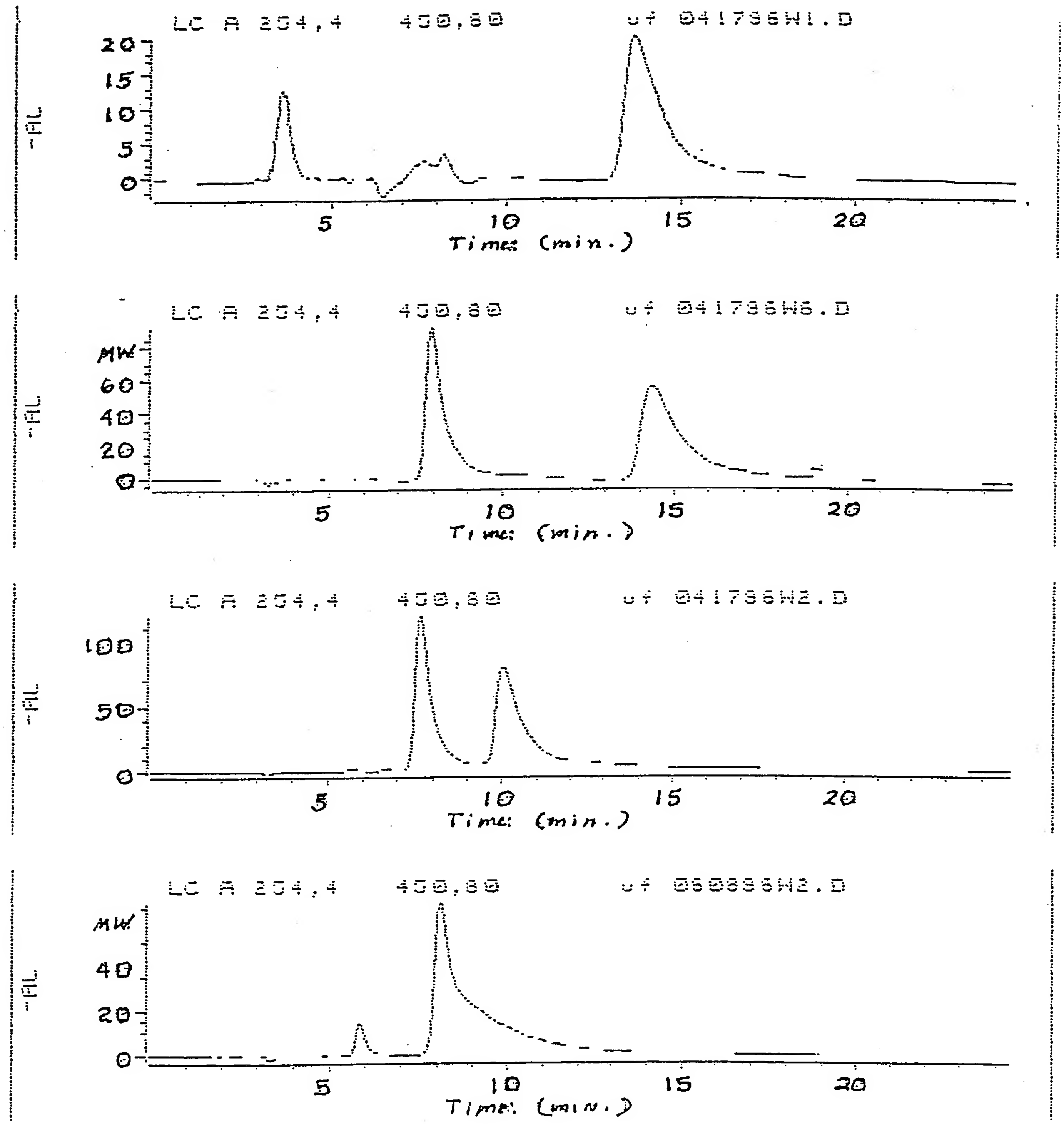

LC A 254,4 450,80 of 041798H1.D
20
15
10
5
0
5
10
15
20
Time (min.)
LC A 254,4 450,80 of 041798H6.D
MW
60
40
20
0
5
10
15
20
Time (min.)
LC A 254,4 450,80 of 041798H2.D
100
50
0
5
10
15
20
Time (min.)
LC A 254,4 450,80 of 050898H2.D
MW
40
20
0
5
10
15
20
Time (min.)

Fig. 6

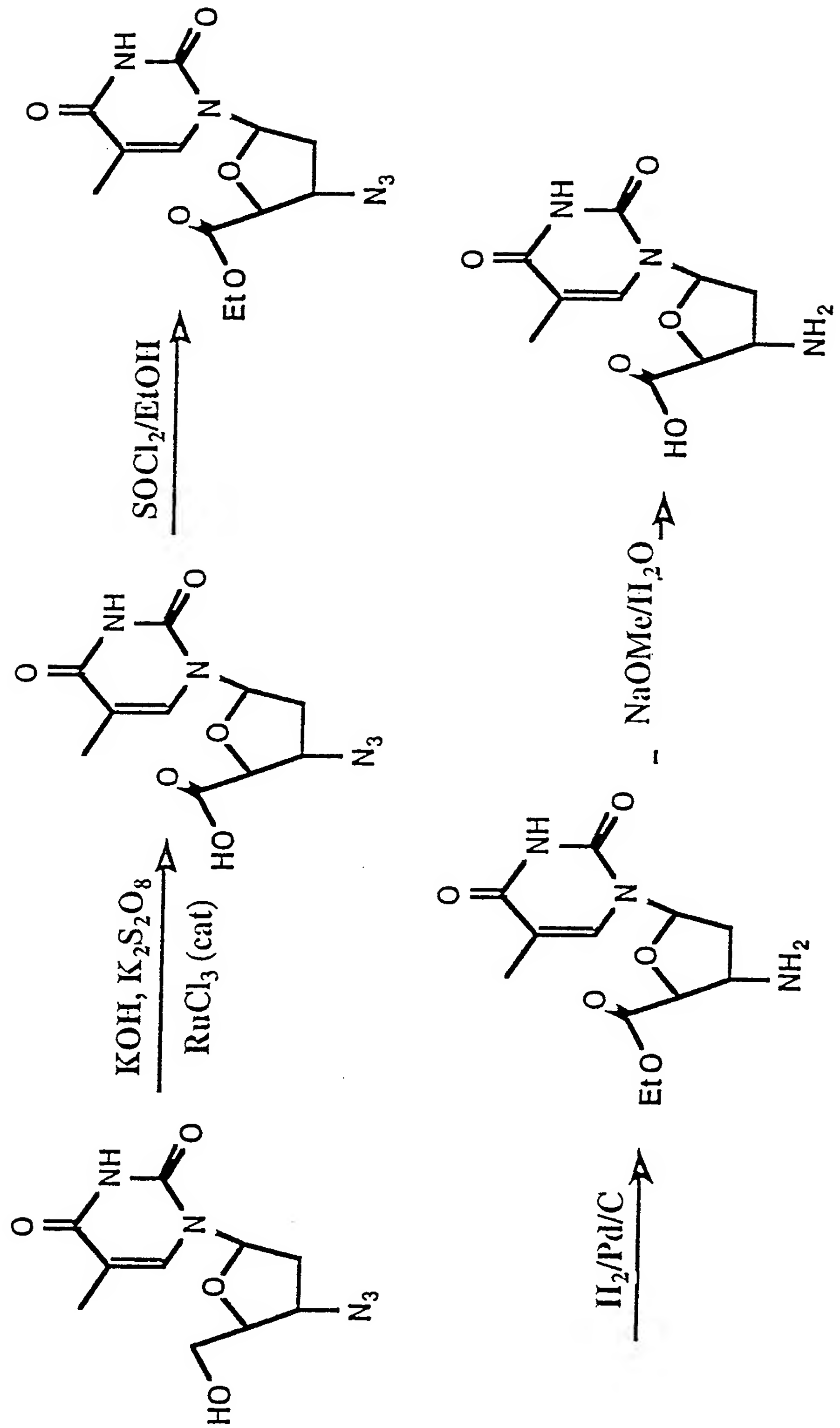


Fig. 7

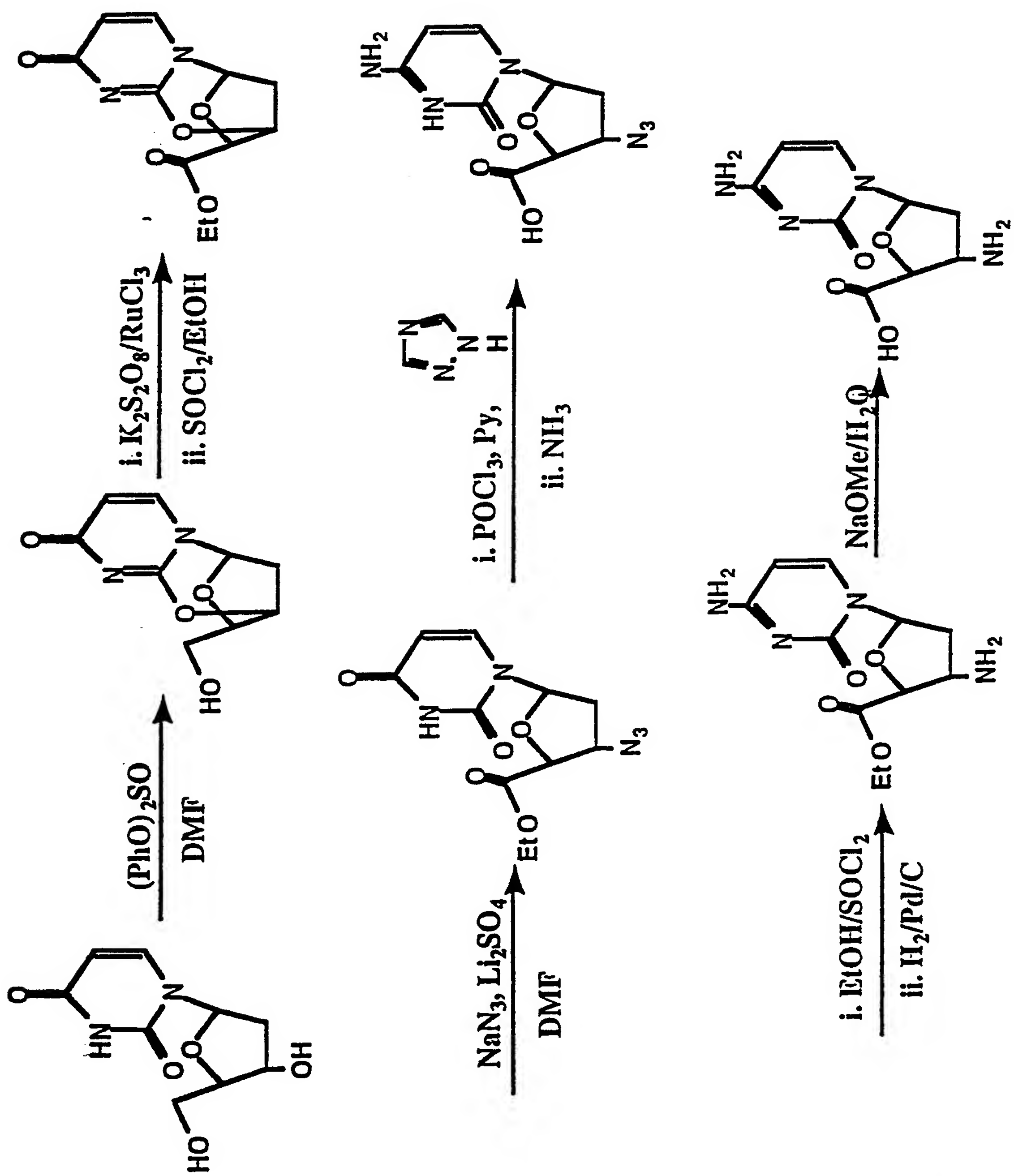

(PhO)$_2$SO
DMF
i. K$_2$S$_2$O$_8$/RuCl$_3$
ii. SOCl$_2$/EtOH
NaN$_3$, Li$_2$SO$_4$
DMF
i. POCl$_3$, Py,
ii. NH$_3$
i. EtOH/SOCl$_2$
ii. H$_2$/Pd/C
NaOMe/H$_2$O

Fig. 8

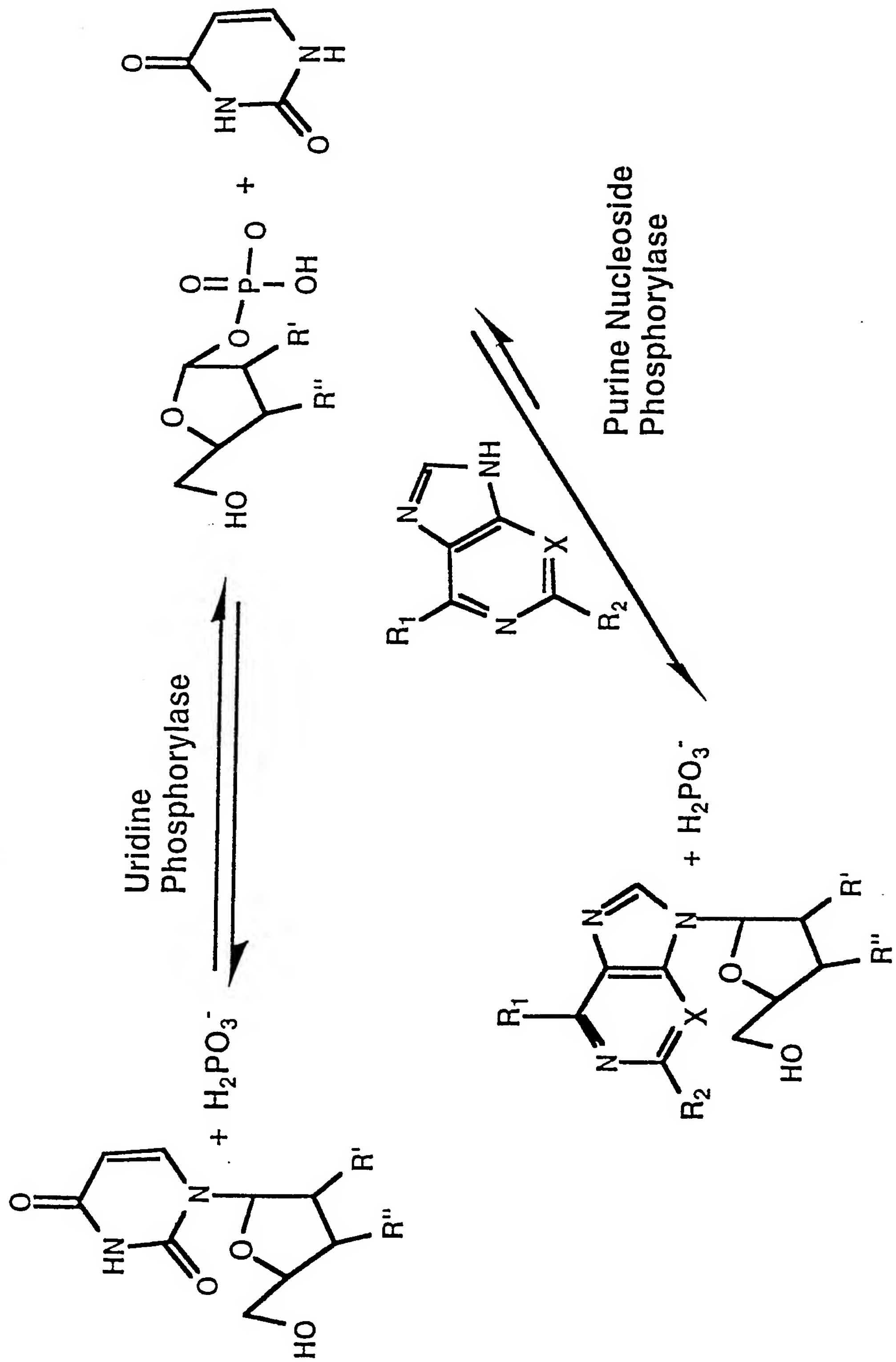
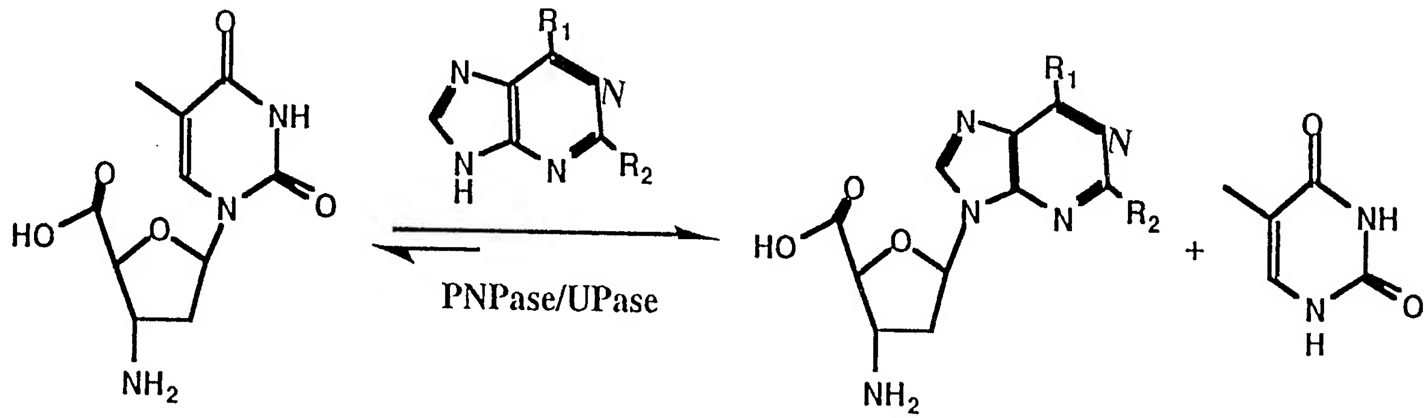


Fig. 9



I, R_1 = NHCbz, R_2 = H
III, R_1 = Cl, R_2 = H

II, R_1 = O, R_2 = NH_2
IV, R_1 = NH_2, R_2 = H

Fig. 10

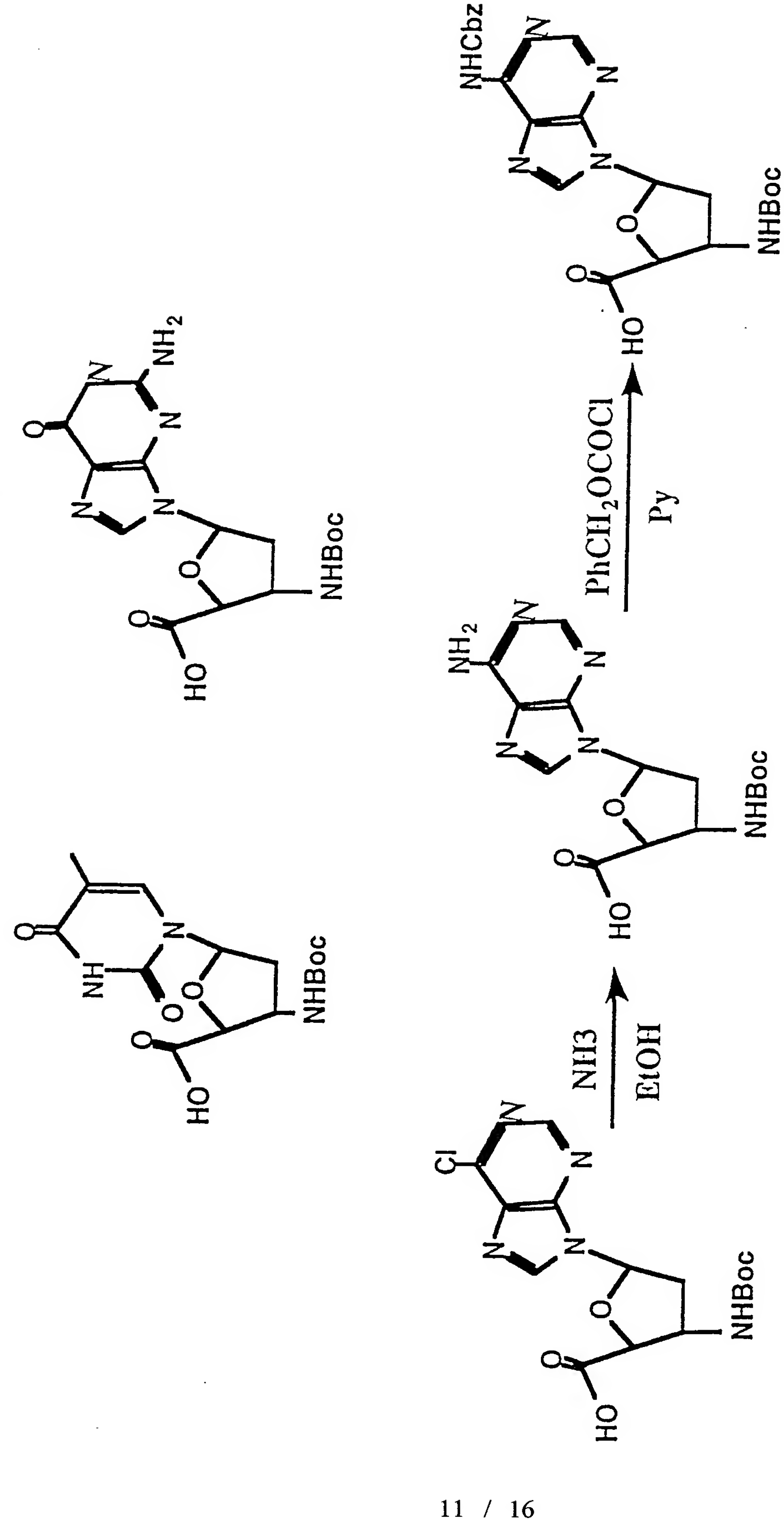


Fig. 11

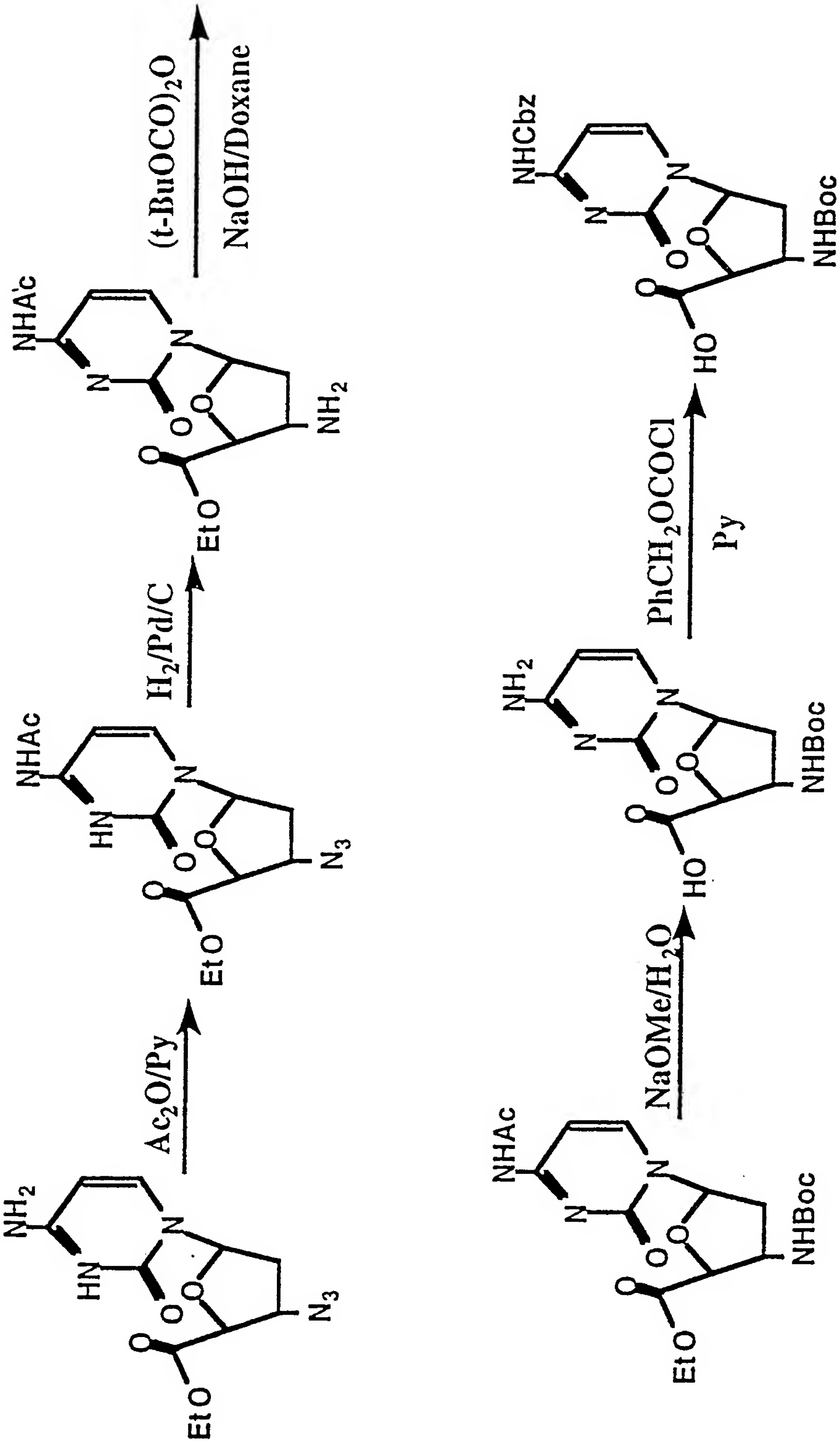

Ac$_2$O/Py
H$_2$/Pd/C
(t-BuOCO)$_2$O
NaOH/Doxane
NaOMe/H$_2$O
PhCH$_2$OCOCl
Py
NH$_2$
NHAc
NHBoc
NHCbz
N$_3$
EtO
HO

Fig. 12

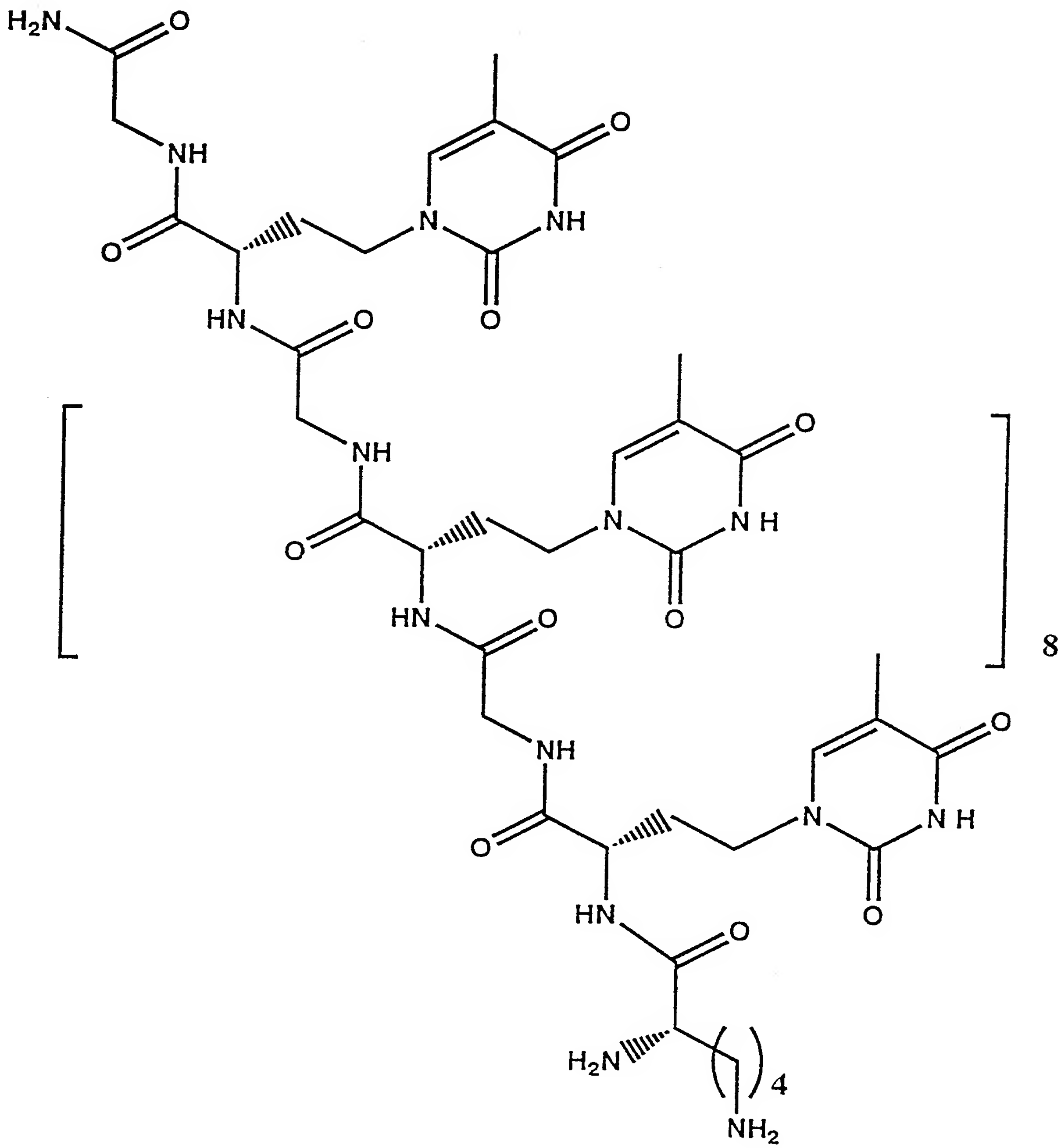


Fig. 13

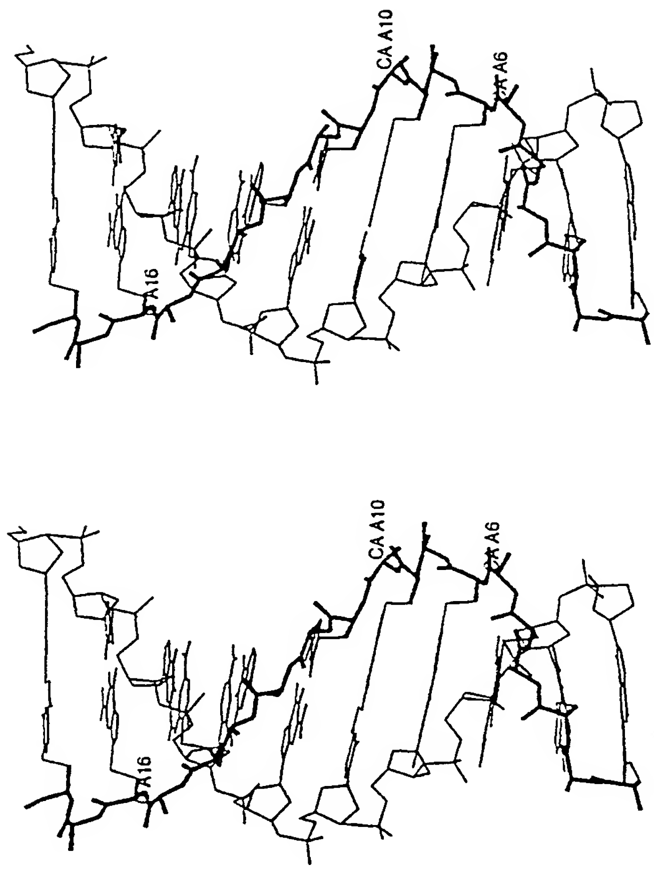


Fig. 14

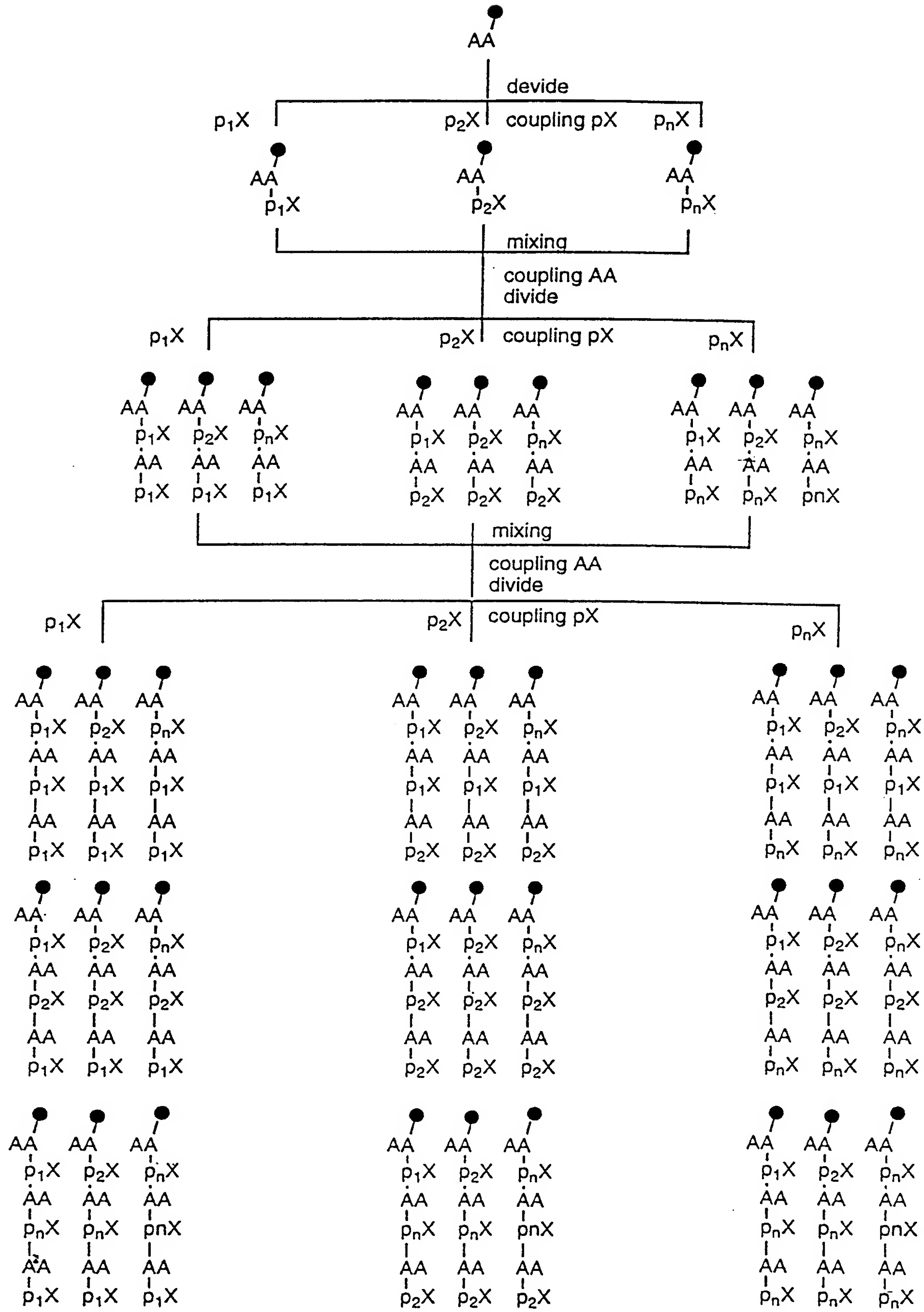


Fig. 15

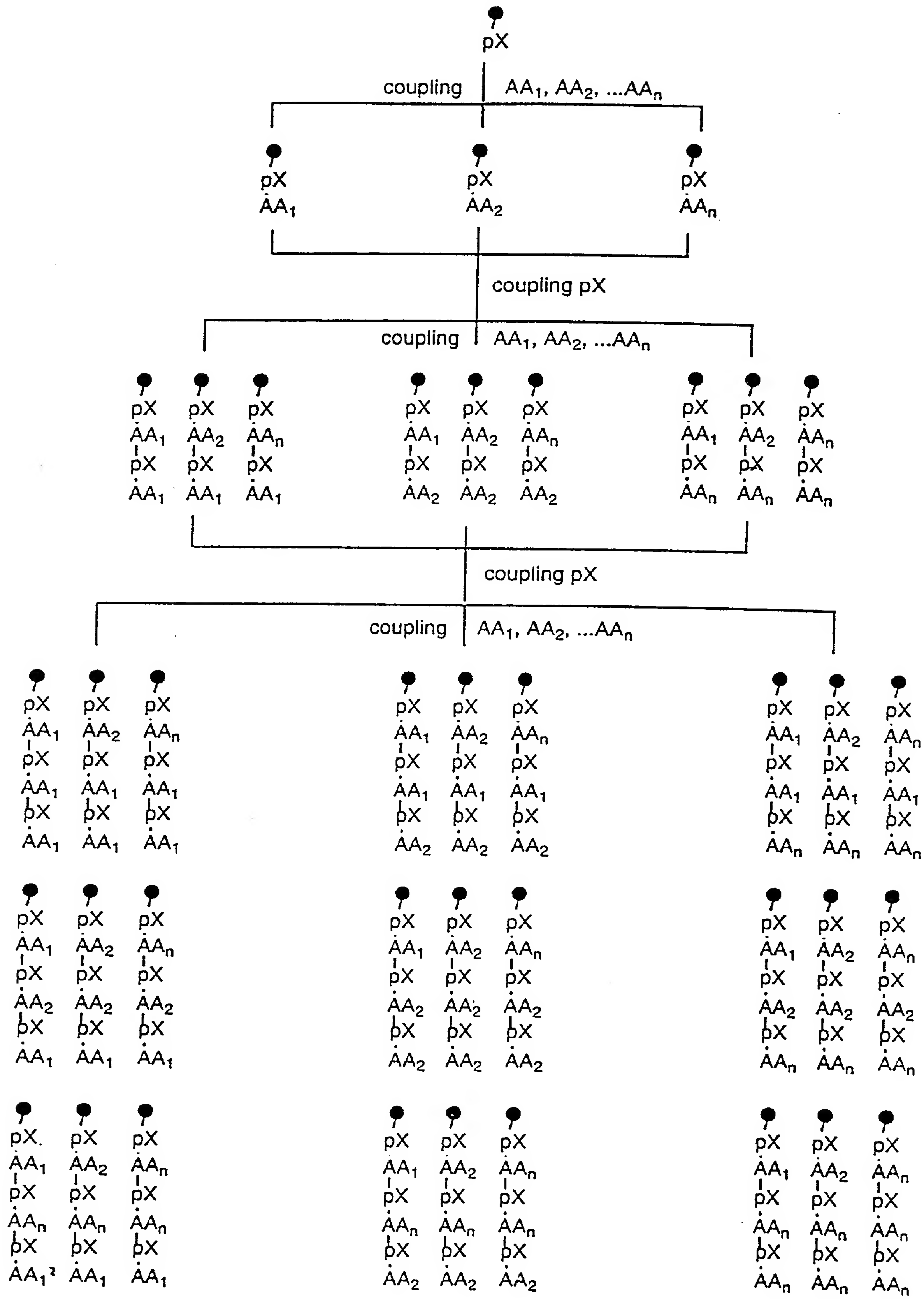

pX
coupling
$AA_1, AA_2, \ldots AA_n$
coupling pX
coupling
$AA_1, AA_2, \ldots AA_n$
coupling pX
coupling
$AA_1, AA_2, \ldots AA_n$

Fig. 16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/12580

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G 01 N 33/53, 33/566, 33/543; C 07 K 1/00, 2/00
US CL : 530/300; 435/7.1; 436/501, 518
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/300; 435/7.1; 436/501, 518

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog
peptide nucleic acids, isomers, chiral, stereoiisomers, diastereoisomers, enantiomers, resolution

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,539,082 A (NIELSEN et al) 23 July 1996 (23/07/06), see entire document, especially figures 2 and columns 1-6.	1-3, 9-12
X	WO 87/04349 A1 (DELLARIA et al) 30 July 1987 (30/07/87), see entire document, especially page 3 and the structures thereon; note that R3 can be heterocyclic ring substituted methyl.	1-3, 9-12 and 15
A	WO 93/12129 A1 (GLAXO Inc.) 24 June 1993 (24/07/93), see entire document especiallythe abstract and the associated figure, note that group "B" can be a purine or pyrimidine as recited on pages 5 and 7.	

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

02 SEPTEMBER 1998

Date of mailing of the international search report

14 OCT 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JOSEPH W. RICIGLIANO

Telephone No. (703) 308-9346

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 92/00091 A1 (Bioligand Inc.) 09 January 1992 (09/01/92), see entire document especially the abstract, and page 12, note that the amino acids encompassed therein include designer amino acids and stereochemically resolved amino acids.	1-3, 9-12 and 15
A	WO 92/20703 A1 (BUCHARDT et al) 26 November 1992 (26,11,92), see entire document especially the incorporation and conjugation of labels on page 14.	